

# Composing short 3-compressing words on a 2 letter alphabet\*

Alessandra Cherubini<sup>1†</sup>Achille Frigeri<sup>1‡</sup>Zuhua Liu<sup>2§</sup><sup>1</sup> Dipartimento di Matematica “F. Brioschi”, Politecnico di Milano<sup>2</sup> Department of Mathematics, Kunming Universityreceived 1998-10-14, revised 17<sup>th</sup> Aug. 2015, accepted tomorrow.

A finite deterministic (semi)automaton  $\mathcal{A} = (Q, \Sigma, \delta)$  is  $k$ -compressible if there is some word  $w \in \Sigma^+$  such that the image of its state set  $Q$  under the natural action of  $w$  is reduced by at least  $k$  states. Such word, if it exists, is called a  $k$ -compressing word for  $\mathcal{A}$ . A word is  $k$ -collapsing if it is  $k$ -compressing for each  $k$ -compressible automaton. We compute a set  $W$  of short words such that each 3-compressible automata on a two letter alphabet is 3-compressed at least by a word in  $W$ . Then we construct a shortest common superstring of the words in  $W$  and, with a further refinement, we obtain a 3-collapsing word of length 53. Moreover, as previously announced, we show that the shortest 3-synchronizing word is not 3-collapsing, illustrating the new bounds  $34 \leq c(2, 3) \leq 53$  for the length  $c(2, 3)$  of the shortest 3-collapsing word on a two letter alphabet.

**Keywords:** deterministic finite automaton, collapsing word, synchronizing word

## 1 Introduction

Let  $\mathcal{A} = (Q, \Sigma, \delta)$  be a finite deterministic complete (semi)automaton with state set  $Q$ , input alphabet  $\Sigma$ , and transition function  $\delta : Q \times \Sigma \rightarrow Q$ . For any word  $w \in \Sigma^+$ , the deficiency of  $w$  is the difference between the cardinality of  $Q$  and the cardinality of the image of  $Q$  under the natural action of  $w$ . For a fixed  $k \geq 1$ , the word  $w$  is called  $k$ -compressing for  $\mathcal{A}$  if its deficiency is greater or equal to  $k$ . An automaton  $\mathcal{A}$  is  $k$ -compressible, if there exists a  $k$ -compressing word for  $\mathcal{A}$ . A word  $w \in \Sigma^+$  is  $k$ -collapsing, if it is  $k$ -compressing for every  $k$ -compressible automaton with input alphabet  $\Sigma$ . A word  $w \in \Sigma^+$  is called  $k$ -synchronizing if it is  $k$ -compressing for all  $k$ -compressible automata with  $k + 1$  states and input alphabet  $\Sigma$ . Obviously each  $k$ -collapsing word is also  $k$ -synchronizing.

The concept of  $k$ -collapsing words arose (under a different name) in the beginning of the 1990s with original motivations coming from combinatorics (Sauer and Stone (1991)) and from abstract algebra

\*The first two authors are partially supported by PRIN: “Automi e linguaggi formali: aspetti matematici e applicativi”. The work was completed during the third-named author’s visit to Politecnico di Milano which was supported by the CSC (China Scholarship Council), and acknowledges support from National Natural Science Foundation of China, grant No. 11261066.

<sup>†</sup>Email: alessandra.cherubini@polimi.it

<sup>‡</sup>Email: achille.frigeri@polimi.it

<sup>§</sup>Email: liuzuhua@hotmail.com

(Pöschel et al. (1994)). In Sauer and Stone (1991) it has been proved that  $k$ -collapsing words always exist, for any  $\Sigma$  and any  $k \geq 1$ , by means of a recursive construction which gives a  $k$ -collapsing word whose length is  $O(2^{2^k})$ . Better bounds for  $c(k, t)$  and  $s(k, t)$ , the length of the shortest  $k$ -collapsing and  $k$ -synchronizing words respectively, on an alphabet of cardinality  $t$  were given in Margolis et al. (2004). The bounds for  $c(2, t)$  were slightly improved in Pribavkina (2005) and Cherubini et al. (2009), but the gaps from lower and upper bounds are quite large even for small values of  $k$  and  $t$ . Exact values of  $s(k, t)$  and  $c(k, t)$  are known for  $k = 2$  and  $t = 2, 3$  and are quite far from the theoretical upper bounds (Ananichev et al. (2005)). Moreover it is known that  $s(3, 2) = 33$  and that the words  $s_{3,2} = ab^2aba^3b^2a^2babab^2a^2b^3aba^2ba^2b^2a$  and its dual  $\bar{s}_{3,2}$  are the unique shortest synchronizing words on  $\{a, b\}$  (Ananichev and Petrov (2003)). Observing that  $s(k, t) \leq c(k, t)$ , and applying the construction in Margolis et al. (2004), one gets  $33 \leq c(3, 2) \leq 154$ .

The reader is referred to Ananichev et al. (2005); Cherubini (2007); Cherubini et al. (2009); Margolis et al. (2004) for references and connections to Theoretical Computer Science and Language Theory. The paper is organized as follows: in Section 2 we introduce some general concepts about 3-compressible automata and the main tool we use to study them, *i.e.*, 3-Missing State Automata. In Section 3 we give a complete characterization of proper 3-compressible automata on a two letter alphabet with a letter acting as a permutation, while in Section 4 we characterize all proper 3-compressible automata without permutations. In Section 5 we describe how to use the previous characterization to obtain a short 3-collapsing word, improving the known upper bound for  $c(3, 2)$ , as already announced in Cherubini et al. (2011). Section 6 ends the paper with some considerations about the quest for short 3-collapsing words in general and the relationship between 3-synchronizing and 3-collapsing words, and how our analysis can be exploited to obtain more general results, as already done in Cherubini and Kisielewicz (2014, 2015).

## 2 Background

Let  $\mathcal{A} = (Q, \Sigma, \delta)$  be a finite deterministic complete (semi)automaton with state set  $Q$ , input alphabet  $\Sigma = \{a, b\}$ , and transition function  $\delta : Q \times \Sigma \rightarrow Q$ . The action of  $\Sigma$  on  $Q$  given by  $\delta$  extends naturally, by composition, to the action of any word  $w \in \Sigma^+$  on  $q \in Q$ ; we denote it by  $qw = \delta(q, w)$ , while the action of  $w$  on the entire state set  $Q$  is denoted by  $Qw = \{qw | q \in Q\}$ .

**Definition 1** *The difference  $|Q| - |Qw|$  is called the deficiency of the word  $w$  with respect to  $\mathcal{A}$  and denoted by  $df_{\mathcal{A}}(w)$ . For a fixed  $k \geq 1$ , a word  $w \in \Sigma^+$  is called  $k$ -compressing for  $\mathcal{A}$ , if  $df_{\mathcal{A}}(w) \geq k$ . An automaton  $\mathcal{A}$  is  $k$ -compressible, if there exists a  $k$ -compressing word for  $\mathcal{A}$ . A word  $w \in \Sigma^+$  is  $k$ -collapsing, if it is  $k$ -compressing for every  $k$ -compressible automaton with input alphabet  $\Sigma$ . A word  $w$  is called  $k$ -synchronizing if it is  $k$ -compressing for all  $k$ -compressible automata with  $k + 1$  states and input alphabet  $\Sigma$ . Obviously each  $k$ -collapsing word is also  $k$ -synchronizing.*

Actually, we view the automaton  $\mathcal{A}$  as a set of *transformations* on  $Q$  induced via  $\delta$  and labeled by letters of  $\Sigma$ , rather than as a standard triple. Indeed, in order to define an automaton, it is enough just to assign to every letter  $a \in \Sigma$  the corresponding transformation  $\tau_a : q \rightarrow \delta(q, a)$  on  $Q$ . Now, for  $a \in \Sigma$ , we get  $df_{\mathcal{A}}(a) = |Q| - |\text{Im}(\tau_a)|$ , hence  $df_{\mathcal{A}}(a) = 0$  if and only if  $\tau_a$  is a permutation on  $Q$ . If  $df_{\mathcal{A}}(a) = m \geq 1$ , then there are exactly  $m$  different states  $y_1, y_2, \dots, y_m \notin \text{Im}(\alpha)$ , and there are some elements of  $Q$  whose images under  $\tau_a$  are equal.

**Definition 2** *Let  $\mathcal{P} = \{\{x_{1_1}, \dots, x_{j_1}\}, \dots, \{x_{1_r}, \dots, x_{j_r}\}\}$  be a partition of  $Q$  (where singleton sets are*

omitted), and  $y_1, \dots, y_m \in Q$ , we say that  $\tau_a$  is a transformation of type

$$[x_{1_1}, \dots, x_{j_1}] \dots [x_{1_r}, \dots, x_{j_r}] \setminus y_1, \dots, y_m$$

if  $\mathcal{P}$  is induced by the kernel of  $\tau_a$  and the states  $y_1, y_2, \dots, y_m$  do not belong to  $\text{Im}(\tau_a)$ .

For instance, if  $\mathcal{A}$  has at least three states denoted by 1, 2 and 3, the transformation  $\tau$  is of type  $[1, 2] \setminus 3$ , if and only if  $\tau(1) = \tau(2)$ , the preimage of 3 is empty, and for any  $q, q' \notin \{1, 2\}$ ,  $\tau(q) = \tau(q')$  if and only if  $q = q'$ . So, with an abuse of language, we will write  $\tau = [1, 2] \setminus 3$  (actually  $[1, 2] \setminus 3$  is a family of transformations). Then, in the sequel we will identify each letter of the input alphabet with its corresponding transformation.

**Definition 3** Let  $a \in \Sigma$ , we say that  $a$  is a permutation letter if it induces a permutation on the set of states, i.e., it has deficiency 0. We assume that permutations on  $Q$ , viewed as elements of the symmetric group  $S_n$  with  $|Q| = n$ , are written in the factorization in disjoint cycles where sometimes also cycles of length 1 are explicitly written. So we will write  $a = (1)(23)\pi$  to denote that (the permutation induces by)  $a$  fixes the state 1, swaps states 2 and 3, and  $\pi$  is a permutation that does not act over  $\{1, 2, 3\}$  ( $\pi$  is not necessarily a cycle).

The notion of transformation induced by a letter naturally extends to words, and then the semigroup generated by the transformations of  $\mathcal{A}$  consists precisely of the transformations corresponding to words in  $\Sigma^+$ . If  $\mathcal{A}$  is  $k$ -compressible, at least one letter of its input alphabet has deficiency greater than 0. It is well known that each  $k$ -collapsing word over a fixed alphabet  $\Sigma$  is  $k$ -full (Sauer and Stone (1991)), i.e., contains each word of length  $k$  on the alphabet  $\Sigma$  among its factors. Hence, to characterize  $k$ -collapsing words it is enough to consider  $k$ -full words compressing all proper  $k$ -compressible automata, i.e.,  $k$ -compressible automata which are not compressed by any word of length  $k$ .

**Proposition 4** Let  $\mathcal{A}$  be a finite complete automaton on the alphabet  $\{a, b\}$ : it is 3-compressible and not proper if at least one letter, say  $a$ , fulfills one of the following conditions:

1. it has deficiency greater than 2;
2. it has deficiency 2 and is of type  $[x, y, z] \setminus u, v$ , with  $\{u, v\} \not\subseteq \{x, y, z\}$ ;
3. it has deficiency 2 and is of type  $[x, y][z, v] \setminus u, w$ , with either  $\{u, w\} = \{x, y\}$ , or  $\{u, w\} = \{z, v\}$ , or  $\{u, w\} \not\subseteq \{x, y, z, v\}$ ;
4. it has deficiency 1 and is of type  $[x, y] \setminus z$ , with  $z \notin \{x, y\}$  and  $za \neq x$ .

The proof of the previous proposition is trivial, indeed if the letter  $a$  fulfills one of the above conditions, then either  $a$  or  $a^2$  or  $a^3$  has deficiency 3. Since we are looking for a 3-compressible proper automaton we may assume that each letter of the alphabet either is a permutation or is of one of the following types (we assume different letters represent different states):

1.  $[x, y, z] \setminus x, y$ ;
2.  $[x, y][z, v] \setminus x, z$ ;
3.  $[x, y] \setminus x$ ;

4.  $[x, y] \setminus z$  with  $za = x$ .

In the sequel we view the set  $Q$  of the states of  $\mathcal{A}$  as a set of natural numbers:  $Q = \{1, 2, \dots, n\}$ , so that, when no confusion arises, a letter  $a$  of types **1**, **2**, **3**, **4** is denoted respectively by  $[1, 2, 3] \setminus 1, 2, [1, 2][3, 4] \setminus 1, 3, [1, 2] \setminus 1, [1, 2] \setminus 3$  with  $3a = 1$ .

**Definition 5** Let  $w \in \Sigma^+$ , we call  $\mathcal{M}(w) = Q \setminus Qw$  the missing set of  $w$ . Let  $Q_1 \subseteq Q$ , we denote by  $\mathcal{M}(Q_1, w)$  the set  $\mathcal{M}(w) \cup \{qw \mid q \in Q_1 \text{ and } \forall q' \in Q \setminus Q_1, qw \neq q'w\}$ , i.e., the missing set of  $w$  having already missed  $Q_1$ .

Observe that  $\mathcal{M}(\emptyset, w) = \mathcal{M}(w)$ , and if  $a \in \Sigma$  is a permutation,  $\mathcal{M}(Q_1, a) = Q_1a$ . Moreover,  $|\mathcal{M}(w)| \geq |\mathcal{M}(w_1)|$ , whenever  $w_1$  is a factor of  $w$ .

**Definition 6** With abuse of language, for a letter  $a$  and  $Q_1 \subseteq Q$ , we call the orbit of  $a$  over  $Q_1$  the set  $\text{Orb}_a(Q_1) = \bigcup_{n=0}^{+\infty} Q_1a^n = \bigcup_{q \in Q_1} \text{Orb}_a(q)$ .

We say that  $\mathcal{A}$  is a  $(\mathbf{i}, \mathbf{j})$ -automaton,  $1 \leq i, j \leq 4$ , if it is an automaton on a two letter alphabet  $\{a, b\}$  and the letter  $a$  is of type  $\mathbf{i}$  and  $b$  is of type  $\mathbf{j}$ . We say that  $\mathcal{A}$  is a  $(\mathbf{i}, \mathbf{p})$ -automaton, with  $1 \leq i \leq 4$ , to denote that the letter  $a$  is of type  $\mathbf{i}$  and  $b$  is a permutation. In the sequel, without loss of generality, we will always suppose that in a  $(\mathbf{i}, \mathbf{j})$ -automaton (resp.  $(\mathbf{i}, \mathbf{p})$ -automaton) the letter  $a$  is of type  $\mathbf{i}$  and  $b$  is of type  $\mathbf{j}$  (resp. a permutation).

Although the notion of missing set is sufficient to describe the compressibility of an automaton, it is in general quite intricate to use, especially when long words are involved. So, to easily calculate the set  $\mathcal{M}(Q_1, w)$ , we introduce a graphical device, the *missing state automaton* of  $\mathcal{A}$ .

**Definition 7** Let  $\mathcal{A} = (Q, \Sigma, \delta)$  be a deterministic (semi)automaton, with  $|Q| = n$  and  $m < n$ . The (complete)  $m$ -Missing State Automaton (*mMSA* for short) of  $\mathcal{A}$  is the automaton  $\mathcal{M} = (\wp^{m-1}(Q) \cup \{\mathbf{m}\}, \Sigma, \emptyset, \tau, \mathbf{m})$ , where  $\wp^{m-1}(Q)$  is the set of subsets of  $Q$  of cardinality less or equal to  $m-1$ ,  $\mathbf{m}$  is a special state not belonging to  $\wp^{m-1}(Q)$  graphically denoted by a circle with  $m$  token inside, and  $\tau : \wp^{m-1}(Q) \times \Sigma \rightarrow \wp^{m-1}(Q) \cup \{\mathbf{m}\}$  is the transition relation defined by

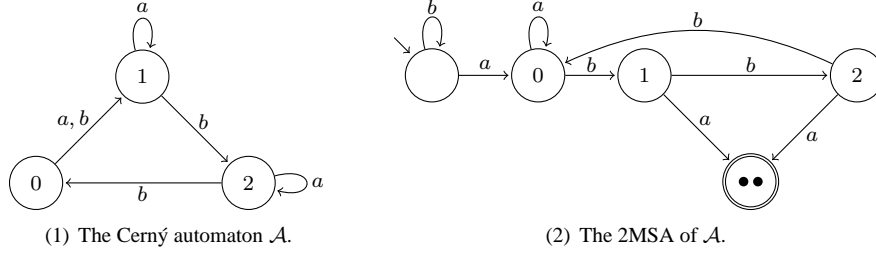
$$\tau(Q_1, a) = \begin{cases} \mathcal{M}(Q_1, a), & \text{if } |\mathcal{M}(Q_1, a)| < m; \\ \mathbf{m}, & \text{else.} \end{cases}$$

notice that  $\tau$  is not defined over  $\mathbf{m}$ .

For example, in Fig. 1 we draw the 2MSA of a simple semiautomaton, proving that it is synchronizable.

The notion of missing state automaton is similar to that of power state automaton, which is a standard tool in computing synchronizing words, see Trahtman (2006); Kudlacik et al. (2012); Volkov (2008), the difference is that the names of states are replaced by their complements and all superstates made by more than  $m$  states are identified. Although, power set automata are only used to design algorithm to find possibly shortest synchronizing words of a fixed automaton, while we need to consider a whole class of automata. Moreover, as we are only interested in knowing if an automaton is 3-compressible, often we will draw only a *Partial 3-Missing State Automaton* (P3MSA), i.e., a path (possibly the shortest) from the initial to a final state of the whole 3MSA.

Lastly, we observe that when considering a family of automata, dozens of subcases arise when we try to find some (short) 3-collapsing word for such family. So, to capture a greater number of cases and improve the readability, we gather several subcases using a *conditional* 3MSA. In such case, a label can



**Fig. 1:** The Cerný semiautomaton with 3 states and its 2MSA: the set of synchronizing words for  $\mathcal{A}$  is the regular language  $b^*a^+(b^3)^*(b+b^2)a(a+b)^*$ .

be of the form  $a|qw \in Q'$ . So,  $\tau(q_1, a|qw \in Q') = q_2$  means that  $\mathcal{M}(q_1, a) = q_2$  under the hypothesis (condition) that  $qw \in Q'$ . Observe that the condition  $qw \in Q'$  spreads over all the states reached by  $q_2$ , so two different states can share the same name, when belonging to different branches. For example, in the conditional 3MSA in Fig. 4(1), the two states named by  $\{1, 3\}$  have different behavior as the one in the first row inherits the condition  $3a = 3$  and then  $\mathcal{M}(1, 3, a) = \{1, 3\}$ , while the one in the second row inherits the condition  $3a = 2$  and then  $\mathcal{M}(1, 3, a) = \{1, 2\}$ .

### 3 3-compressible (i, p)-automata

In this section we characterize all proper 3-compressible automata over the alphabet  $\{a, b\}$  in which the letter  $b$  acts as a permutation on the set  $Q$  of states. In particular in the following propositions we give a small set of short 3-collapsing words when letter  $a$  is of type **i**,  $1 \leq i \leq 4$ .

**Proposition 8** *Let  $\mathcal{A}$  be a (1, p)-automaton with  $a = [1, 2, 3] \setminus 1, 2$ . Then  $\mathcal{A}$  is 3-compressible and proper if, and only if, the following conditions hold:*

1.  $\text{Orb}_b(1, 2) \not\subseteq \{1, 2, 3\}$ , and
2.  $\{1, 2\}b \subseteq \{1, 2, 3\}$ .

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then the word  $ab^2a$  3-compresses  $\mathcal{A}$ .

**Proof:** Let  $\mathcal{A}$  be a (1, p)-automaton that does not satisfy one of the conditions 1. and 2. If  $\text{Orb}_b(1, 2) \subseteq \{1, 2, 3\}$  for all word  $w \in \{a, b\}^*$  it is  $\mathcal{M}(wa) = \{1, 2\}$ , then  $\mathcal{A}$  is not 3-compressible. Else, if  $\{1, 2\}b \not\subseteq \{1, 2, 3\}$ , then  $\mathcal{M}(a) = \{1, 2\}$ ,  $\mathcal{M}(ab) = \{1b, 2b\} \not\subseteq \{1, 2, 3\}$ , and  $|\mathcal{M}(aba)| = 3$ , so  $\mathcal{A}$  is not proper.

Conversely, let  $\mathcal{A}$  be an automaton satisfying conditions 1. and 2. A 3-compressing word for  $\mathcal{A}$  must have at least two non-consecutive occurrences of letter  $a$ , and the word  $aba$  is not 3-compressing. Moreover,  $\{1, 2\}b^2 \not\subseteq \{1, 2, 3\}$ , else  $\text{Orb}_b(1, 2) \subseteq \{1, 2, 3\}$ , against the hypothesis, and then the word  $ab^2a$  3-compresses  $\mathcal{A}$ .  $\square$

**Proposition 9** *Let  $\mathcal{A}$  be a (2, p)-automaton with  $a = [1, 2][3, 4] \setminus 1, 3$ . Then  $\mathcal{A}$  is 3-compressible and proper if, and only if,*

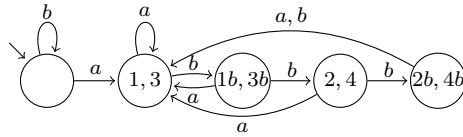
1.  $\text{Orb}_b(1, 3) \not\subseteq \{1, 2, 3, 4\}$  and  $\{1, 3\}b \subseteq \{1, 2, 3, 4\}$ , or

2.  $\{1, 3\}b \in \{\{1, 4\}, \{2, 3\}\}$  and either  $|Orb_b(1)| = 3$  or  $|Orb_b(3)| = 3$ .

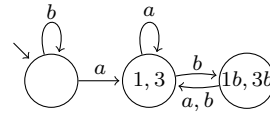
Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then one of the words  $ab^2a$  or  $ab^3a$  3-compresses  $\mathcal{A}$ .

**Proof:** Let  $\mathcal{A}$  a  $(2, \mathbf{p})$ -automaton that does not satisfy both conditions 1. and 2. If  $\{1, 3\}b \not\subseteq \{1, 2, 3, 4\}$ , then the word  $aba$  3-compresses  $\mathcal{A}$ , which is not proper. So, let  $Orb_b(1, 3) \subseteq \{1, 2, 3, 4\}$ , then we have to consider only the following cases:

1.  $\{1, 3\}b \in \{\{1, 2\}, \{3, 4\}\}$ , then again the word  $aba$  3-compresses  $\mathcal{A}$ ;
2.  $\{1, 3\}b = \{1, 3\}$ , then for all  $w \in b^*$ ,  $\mathcal{M}(w) = \emptyset$ , while for all  $w \in \{a, b\}^* \setminus b^*$ ,  $\mathcal{M}(w) = \{1, 3\}$ , then  $\mathcal{A}$  is not 3-compressible;
3.  $\{1, 3\}b \in \{\{1, 4\}, \{2, 3\}\}$  with  $|Orb_b(1)| \neq 3$  or  $|Orb_b(3)| \neq 3$ , then  $b = (1423)\pi$  or  $b = (1324)\pi$  or  $b = (1)(2)(34)\pi$  or  $b = (12)(3)(4)\pi$ . The 3MSA in figures 2(1) and 2(2) prove that in any case  $\mathcal{A}$  is not 3-compressible;
4.  $\{1, 3\}b = \{2, 4\}$ , in this case we have four subcases:  $b = (12)(34)\pi$  or  $b = (14)(32)\pi$  or  $b = (1234)\pi$  or  $b = (1432)\pi$ , and the 3MSA in Fig. 2(2) proves that  $\mathcal{A}$  is not 3-compressible.



(1) 3MSA for the case  $b = (1423)\pi$  or  $b = (1324)\pi$ .



(2) 3MSA for the case  $b = (1)(2)(34)\pi$ ,  $b = (12)(3)(4)\pi$ , or  $\{1, 3\}b = \{2, 4\}$ .

**Fig. 2:** 3MSA for automata that do not satisfy conditions 1. and 2. of Proposition 9.

Conversely,

1. if  $Orb_b(1, 3) \not\subseteq \{1, 2, 3, 4\}$  and  $\{1, 3\}b \subseteq \{1, 2, 3, 4\}$ . Then  $\mathcal{A}$  is proper and either  $1b \neq 1$  or  $3b \neq 3$ , and so  $|\{1, 3, 1b, 3b\}| \geq 3$ . If  $|\{1, 3, 1b, 3b\}| = 4$ , then  $|\{1, 3, 1b, 3b, 1b^2, 3b^2\}| > 4$  and the word  $ab^2a$  3-compresses  $\mathcal{A}$ . Otherwise, let  $|\{1, 3, 1b, 3b\}| = 3$ . Without loss of generality, we can assume  $1b = 1$  and  $3b \neq 3$ . As  $3b \in \{2, 4\}$ ,  $\mathcal{A}$  is proper. If  $3b^2 \notin \{1, 2, 3, 4\}$ , then the word  $ab^2a$  3-compresses  $\mathcal{A}$ . If  $3b^2 \in \{1, 2, 3, 4\}$ , actually  $3b^2 \in \{2, 4\}$ . As  $3b \neq 3b^2$ , then  $|\{1, 3, 3b, 3b^2\}| = 4$ , so  $3b^3 \notin \{1, 2, 3, 4\}$  and the word  $ab^3a$  3-compresses  $\mathcal{A}$ .
2. Suppose  $\{1, 3\}b = \{1, 4\}$ . If  $|Orb_b(1)| = 3$ , then  $1b = 4$  and  $3b = 1$ , i.e.,  $b = (143)\pi$ , if  $|Orb_b(3)| = 3$ , then either  $b = (143)\pi$ , or  $b$  fixes 1 and is of the form  $(34x)\pi$  for some  $x \notin \{1, 3, 4\}$ . Then  $\{1, 3\}b^2 \in \{\{1, x\}, \{3, 4\}\}$ , so the word  $ab^2a$  3-compresses  $\mathcal{A}$ . The case  $\{1, 3\}b = \{2, 3\}$  is identical.

□

Observe that each 3-compressible  $(3, \mathbf{p})$ - or  $(4, \mathbf{p})$ -automaton  $\mathcal{A}$  is proper. Indeed each 3-compressing word for  $\mathcal{A}$  contains at least three occurrences of the letter  $a$  which are not all consecutive. So in the next two propositions, we only look for 3-compressible automata.

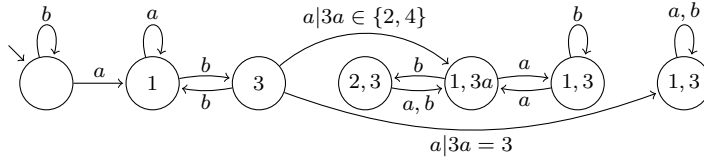
**Proposition 10** *Let  $\mathcal{A}$  be a  $(3, \mathbf{p})$ -automaton with  $a = [1, 2] \setminus 1$ . Then  $\mathcal{A}$  is 3-compressible (and proper) if, and only if, the following conditions hold:*

1.  $|Orb_b(1)| \geq 2$  and  $\{1, 2\}b \neq \{1, 2\}$ , and
2. if  $b = (13)\pi$ , then
  - (a)  $3a \neq 3$ , and
  - (b) if  $3a = 2$ , then  $2b \neq 2$  or  $2a \neq 3$ , and
  - (c) if  $3a = 2b$  and  $2b \notin \{2, 3\}$ , then  $2b^2 \neq 2$  and  $2ba \neq 3$ ,
3. if  $b = (123)\pi$ , or  $b = (132)\pi$ , then  $\{2, 3\}a \neq \{2, 3\}$ , and
4. if  $b = (1324)\pi$ , then  $\{3, 4\}a \neq \{3, 4\}$ .

Moreover, if  $\mathcal{A}$  is 3-compressible (and proper), then one of the words  $ababa$  or  $aba^2ba$  or  $ab^2ab^2a$  or  $ab^2a^2b^2a$  or  $ab^2abab^2a$  or  $abab^2aba$  or  $ab^3aba$  or  $abab^3a$  or  $ab^3ab^3a$  3-compresses  $\mathcal{A}$ .

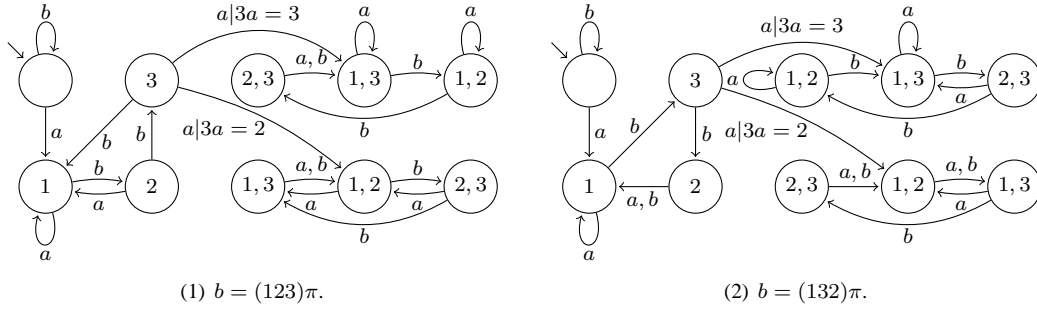
**Proof:** Let  $\mathcal{A}$  be a  $(3, \mathbf{p})$ -automaton that does not satisfy one of the conditions 1.-4., we prove that it is not 3-compressible.

1. Let  $|Orb_b(1)| = 1$  or  $\{1, 2\}b = \{1, 2\}$ , then  $1b = 1$  or  $b = (12)\pi$ . In the former case for all  $w \in (a + b)^*$ ,  $\mathcal{M}(w) \in \{\emptyset, \{1\}\}$ , in the latter  $\mathcal{M}(w) \in \{\emptyset, \{1\}, \{2\}\}$ , so  $\mathcal{A}$  is not 3-compressible.
2. Let  $b = (13)\pi$  and either  $3a = 3$ , or  $3a = 2$ ,  $2b = 2$  and  $2a = 3$ , or  $3a = 2b = 4$ ,  $2b^2 = 2$  and  $2ba = 3$  (in the last two cases  $3a^2 = 3$  and  $3ab = 2$ ). Then the 3MSA in Fig. 3 proves that in any case  $\mathcal{A}$  is not 3-compressible.

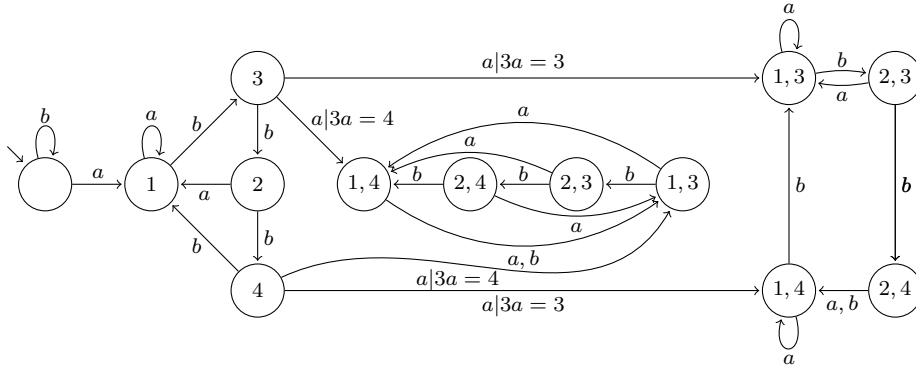


**Fig. 3:** 3MSA for automata that do not satisfy condition 2. of Proposition 10.

3. Let  $b = (123)\pi$ , or  $b = (132)\pi$ , and  $\{2, 3\}a = \{2, 3\}$ . The 3MSA in figures 4(1) and 4(2) prove that  $\mathcal{A}$  is not 3-compressible.
4. Let  $b = (1324)\pi$ , and  $\{3, 4\}a = \{3, 4\}$ . The 3MSA in Fig. 5 proves that  $\mathcal{A}$  is not 3-compressible.



**Fig. 4:** 3MSA for automata that do not satisfy condition 3. of proposition 10.

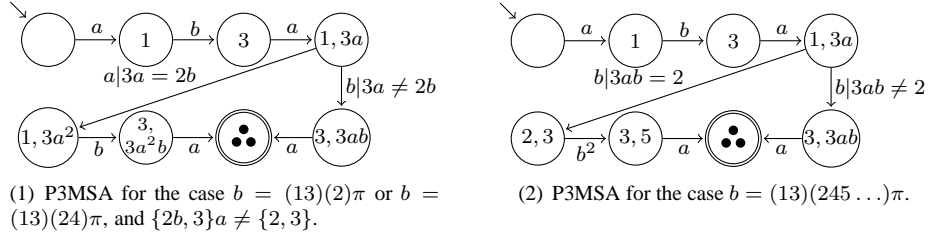


**Fig. 5:** 3MSA for automata that do not satisfy condition 4. of proposition 10.

Conversely, let  $\mathcal{A}$  be an automaton satisfying conditions 1. – 4. Since  $1b \neq 1$  and  $\{1, 2\}b \neq \{1, 2\}$ , then  $|Orb_b(1)| \geq 2$  and  $b$  is not of the form  $(12)\pi$ .

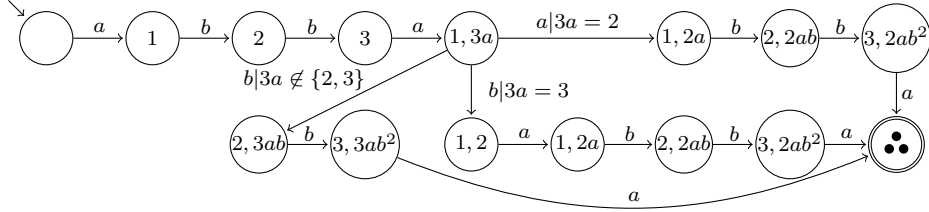
1. Let  $|Orb_b(1)| = 2$ ,  $b = (13)\pi$ , and  $3a \neq 3$ , then there are two further subcases:
  - (a) if  $|Orb_b(2)| \leq 2$ , i.e.,  $b = (13)(2)\pi$  or  $b = (13)(24)\pi$ , then  $\{2b, 3\}a \neq \{2b, 3\}$  and in Fig. 6(1) we draw a P3MSA for those cases, proving that either the word  $ababa$  or  $aba^2ba$  3-compresses  $\mathcal{A}$ .
  - (b) if  $|Orb_b(2)| \geq 3$ , i.e.,  $b = (13)(245\dots)\pi$ , then in Fig. 6(2) we draw a P3MSA for this case, proving that either the word  $ababa$  or  $abab^3a$  3-compresses  $\mathcal{A}$ .
2. Let  $|Orb_b(1)| = 3$ , then there are three further subcases:
  - (a) if  $b = (123)\pi$ , then  $\{2, 3\}a \neq \{2, 3\}$  and in Fig. 7 we draw a P3MSA for this case, proving that either the word  $ab^2ab^2a$  or  $ab^2a^2b^2a$  or  $ab^2abab^2a$  3-compresses  $\mathcal{A}$ .
  - (b) if  $b = (132)\pi$ , then  $\{2, 3\}a \neq \{2, 3\}$  and in Fig. 8 we draw a P3MSA for this case, proving that either the word  $aba^2ba$  or  $ababa$  or  $abab^2aba$  3-compresses  $\mathcal{A}$ .



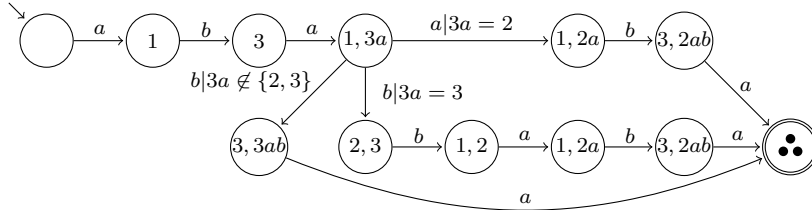


**Fig. 6:** P3MSA for 3-compressible automata with  $b = (13)\pi$  and  $3a \neq 3$ .

- (c) if  $b = (134)\pi$ , then in Fig. 9 we draw a P3MSA for this case, proving that either the word  $abab^2a$  or  $ababa$  or  $ab^2ab^2a$  or  $ab^2aba$  3-compresses  $\mathcal{A}$ .



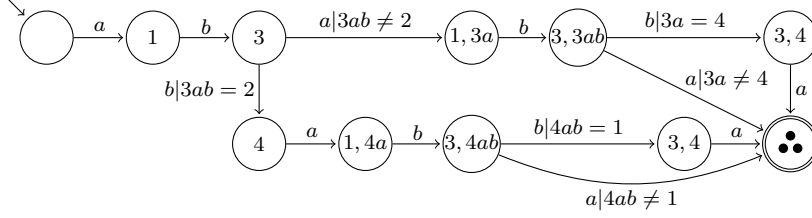
**Fig. 7:** P3MSA for 3-compressible automata with  $b = (123)\pi$  and  $\{2, 3\}a \neq \{2, 3\}$ .



**Fig. 8:** P3MSA for 3-compressible automata with  $b = (132)\pi$  and  $\{2, 3\}a \neq \{2, 3\}$ .

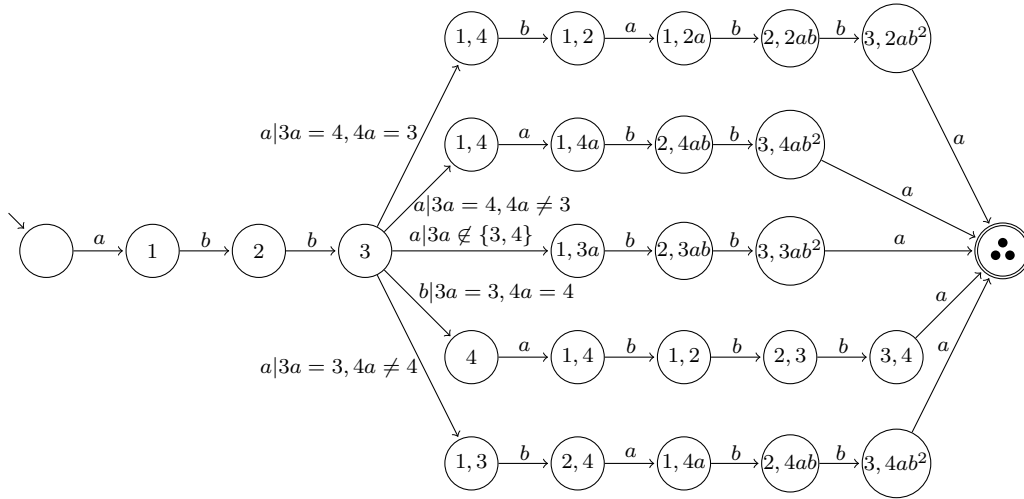
3. Let  $|Orb_b(1)| = 4$ , then there are three further subcases:

- (a) if  $b = (1234)\pi$ , then in Fig. 10 we draw a P3MSA for this case, proving that either the word  $ab^2ab^2a$  or  $ab^3ab^3a$  or  $ab^2abab^2a$  or  $ab^2a^2b^2a$  3-compresses  $\mathcal{A}$ . Indeed, observe that:
- if  $3a \notin \{3, 4\}$ , then  $3ab^2 = 1$  implies  $3a = 3$ , and  $3ab^2 = 2$  implies  $3a = 4$ , and both of them are contradictions;
  - if  $3a = 4$  and  $4a = 3$ , then  $2ab^2 = 1$  implies  $2a = 3$ , and  $2ab^2 = 2$  implies  $2a = 4$ , both contradictions;



**Fig. 9:** P3MSA for 3-compressible automata with  $b = (134)\pi$ .

- iii. if  $3a = 3$  and  $4a \neq 4$  or if  $3a = 4$  and  $4a \neq 3$ , then  $4ab^2 = 1$  implies  $4a = 3$ , and  $4ab^2 = 2$  implies  $4a = 4$ , both contradictions.

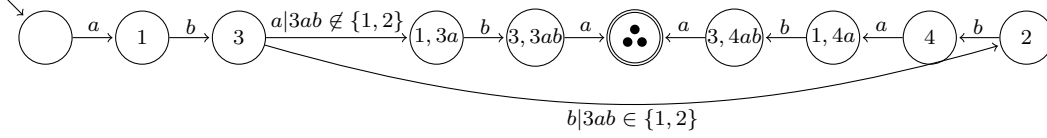


**Fig. 10:** P3MSA for the case  $b = (1234)\pi$ .

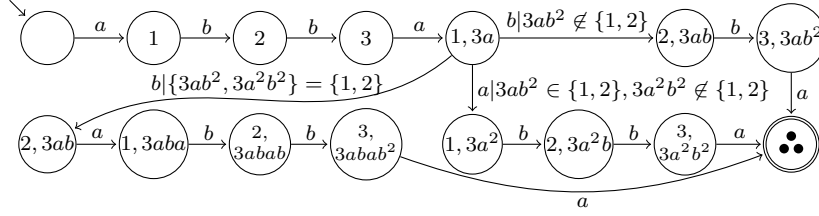
- (b) if  $b = (1324)\pi$ , then  $\{3, 4\}a \neq \{3, 4\}$  and in Fig. 11 we draw a P3MSA for this case, proving that either the word  $ababa$  or  $ab^3aba$  3-compresses  $\mathcal{A}$ . Observe that if  $3ab = 1$ , then  $3a = 4$ ,  $4a \neq 3$ , and so  $4ab \notin \{1, 2\}$ , else if  $3ab = 2$ , then  $3a = 3$ ,  $4a \notin \{3, 4\}$ , and again  $4ab \notin \{1, 2\}$ .

4. Let  $|Orb_b(1)| \geq 5$ , then there are three further subcases:

- (a) if  $b = (12345 \dots)\pi$ , then in Fig. 12 we draw a P3MSA for this case, proving that either the word  $ab^2ab^2a$  or  $ab^2abab^2a$  or  $ab^2a^2b^2a$  3-compresses  $\mathcal{A}$ . Note that if  $3ab^2 = 1$  and  $3a^2b^2 = 2$ , then  $3abab^2 \notin \{1, 2\}$ . In fact if  $3abab^2 = 1$ , then  $3a = 3aba$ , hence  $3a = 2$  and  $3ab^2 = 4$ , a contradiction. If  $3abab^2 = 2$ , then  $3aba = 3a^2$ , hence  $3ab = 3a$  and  $3ab^2 = 3ab = 1$ , again a contradiction. Similarly, if  $3ab^2 = 2$  and  $4ab^3 = 2$ , then  $3abab^2 \notin \{1, 2\}$ ; in fact if  $3abab^2 = 1$  then  $3ab^2ab^2 = 4ab^2$  and  $3ab^2 = 4$ , else if  $3abab^2 = 2$  then  $3ab^2 = 3abab^2$  hence  $3ab = 3$  and  $3ab^2 = 4$ , and in both cases this is a contradiction.

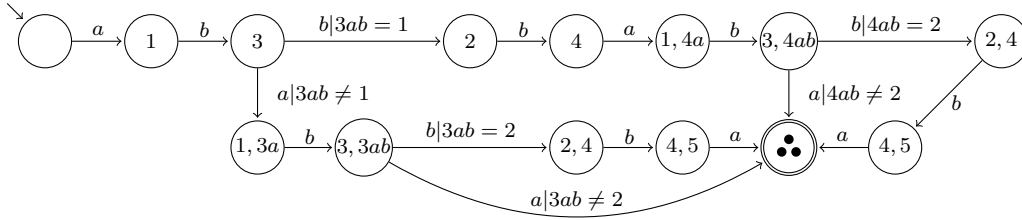


**Fig. 11:** P3MSA for the case  $b = (1324)\pi$  and  $\{3, 4\}a \neq \{3, 4\}$ .



**Fig. 12:** P3MSA for the case  $b = (12345 \dots)\pi$ .

- (b) if  $b = (12345 \dots)\pi$ , then in Fig. 13 we draw a P3MSA for this case, proving that either the word  $ab^3ab^3a$  or  $ab^3aba$  or  $abab^3a$  or  $ababa$  3-compresses  $\mathcal{A}$ .



**Fig. 13:** 3MSA for the case  $b = (12345 \dots)\pi$ .

- (c) if  $b = (13425 \dots)\pi$ , then  $\mathcal{M}(abab) = \{3, 3ab\}$ . If  $3ab \notin \{1, 2\}$ , then the word  $ababa$  3-compresses  $\mathcal{A}$ , else if  $3ab = 1$ , then  $\mathcal{M}(abab^2) = \{3, 4\}$ , else if  $3ab = 2$ , then  $\mathcal{M}(abab^2) = \{4, 5\}$  and in both cases the word  $abab^2a$  3-compresses  $\mathcal{A}$ .

□

**Proposition 11** Let  $\mathcal{A}$  be a  $(4, \mathbf{p})$ -automaton with  $a = [1, 2] \setminus 3, 3a = 1$ . Then  $\mathcal{A}$  is 3-compressible (and proper) if, and only if, the following conditions hold:

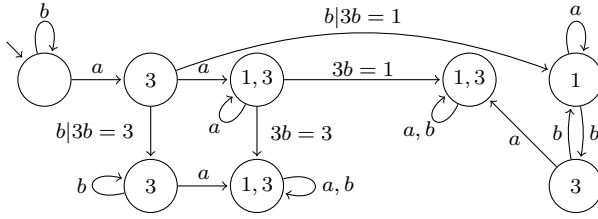
1.  $\{1, 3\}b \neq \{1, 3\}$  and
2.  $b \neq (12)(3)\pi$  and
3. if  $b = (1)(23)\pi$ , or  $b = (123)\pi$ , or  $b = (132)\pi$ , then  $2a \neq 2$  and

4. if  $b = (1)(2)(34)\pi$ , or  $b = (12)(34)\pi$ , or  $b = (14)(23)\pi$ , or  $b = (1423)\pi$ , or  $b = (1324)\pi$ , then  $4a \neq 2$ .

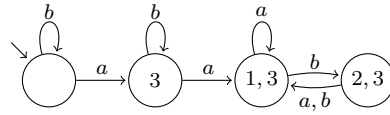
Moreover, if  $\mathcal{A}$  is 3-compressible (and proper), then one of the words  $a^2ba^2$  or  $a^2b^2a^2$  or  $a^2b^3a$  or  $a^2baba^2$  or  $ab^3ab^3a$  3-compresses  $\mathcal{A}$ .

**Proof:** Let  $\mathcal{A}$  be a  $(4, \mathbf{p})$ -automaton that does not satisfy one of the conditions 1. – 4., we prove that it is not 3-compressible.

1. Let condition 1. be false, i.e.,  $\{1, 3\}b = \{1, 3\}$ , then the 3MSA in Fig. 14(1) proves that  $\mathcal{A}$  is not 3-compressible.
2. Let condition 2. be false, i.e.,  $b = (12)(3)\pi$ , then the 3MSA in Fig. 14(2) proves that  $\mathcal{A}$  is not 3-compressible.



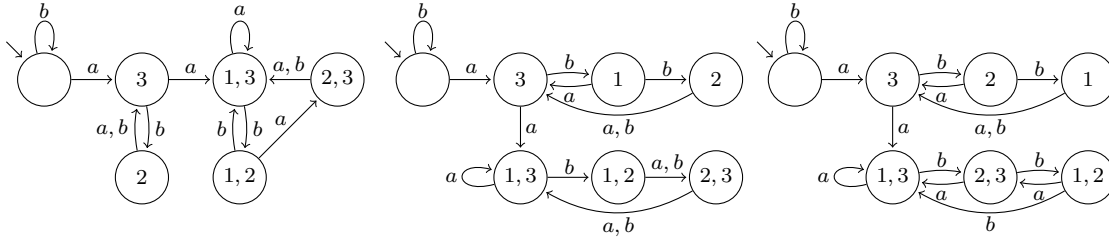
(1) 3MSA for the case  $\{1, 3\}b = \{1, 3\}$ .



(2) 3MSA for the case  $b = (12)(3)\pi$ .

**Fig. 14:** 3MSA for automata that do not satisfy conditions 1. or 2. of Proposition 11.

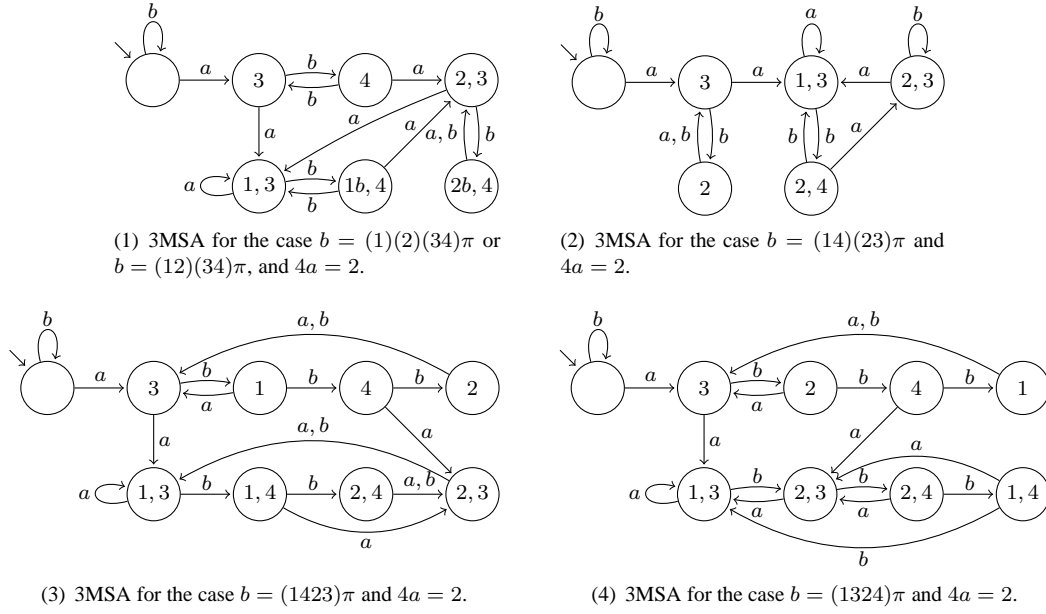
3. Let condition 3. be false, i.e.,  $b = (1)(23)\pi$  or  $b = (123)\pi$  or  $b = (132)\pi$ , and  $2a = 2$ . Then the 3MSA in figures 15(1), 15(2) and 15(3) prove that  $\mathcal{A}$  is not 3-compressible.



(1) 3MSA for the case  $b = (1)(23)\pi$  and  $2a = 2$ . (2) 3MSA for the case  $b = (123)\pi$  and  $2a = 2$ . (3) 3MSA for the case  $b = (132)\pi$  and  $2a = 2$ .

**Fig. 15:** 3MSA for automata that do not satisfy condition 3. of Proposition 11.

4. Let condition 4. be false, i.e.,  $b = (1)(2)(34)\pi$  or  $b = (12)(34)\pi$  or  $b = (14)(23)\pi$  or  $b = (1423)\pi$  or  $b = (1324)\pi$ , and  $4a = 2$ . Then the 3MSA in figures 16(1), 16(2), 16(3) and 16(4) prove that  $\mathcal{A}$  is not 3-compressible.



**Fig. 16:** 3MSA for automata that do not satisfy condition 4. of Proposition 11.

Conversely, let  $\mathcal{A}$  be an automaton satisfying conditions 1. – 4..

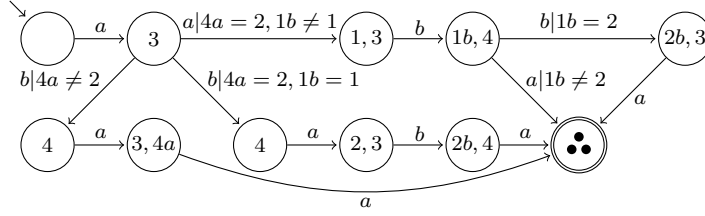
If  $Orb_b(1, 3) \subseteq \{1, 2, 3\}$ , then  $Orb_b(1, 3) = \{1, 2, 3\}$ , as  $\{1, 3\}b \neq \{1, 3\}$ . There are two subcases:

1. if  $b = (1)(23)\pi$  or  $b = (123)\pi$ , and  $2a \neq 2$ , then  $\mathcal{M}(a^2b) = \{1, 2\}$ ,  $\mathcal{M}(a^2ba) = \{2a, 3\}$ , and  $\mathcal{M}(a^2ba^2) = \{1, 3, 2a^2\}$ , and the word  $a^2ba^2$  3-compresses  $\mathcal{A}$ ;
2. if  $b = (132)\pi$  and  $2a \neq 2$ , then  $\mathcal{M}(a^2b) = \{2, 3\}$ ,  $\mathcal{M}(a^2b^2) = \{1, 2\}$ ,  $\mathcal{M}(a^2b^2a) = \{2a, 3\}$ , and  $\mathcal{M}(a^2b^2a^2) = \{1, 3, 2a^2\}$ , and the word  $a^2b^2a^2$  3-compresses  $\mathcal{A}$ .

Let now  $Orb_b(1, 3) \not\subseteq \{1, 2, 3\}$ , we distinguish four subcases by considering the cardinality of  $Orb_b(3)$ .

1. Let  $Orb_b(3) = \{3\}$ , then there are two further subcases:
  - (a) if  $b = (14 \dots)(3)\pi$ , then  $\mathcal{M}(a^2b) = \{3, 4\}$ ,  $\mathcal{M}(a^2ba) = \{1, 3, 4a\}$ , and the word  $a^2ba$  3-compresses  $\mathcal{A}$ ;
  - (b) if  $b = (124 \dots)(3)\pi$ , then  $\mathcal{M}(a^2b) = \{2, 3\}$ ,  $\mathcal{M}(a^2b^2) = \{3, 4\}$ ,  $\mathcal{M}(a^2b^2a) = \{1, 3, 4a\}$ , and the word  $a^2b^2a$  3-compresses  $\mathcal{A}$ .
2. Let  $Orb_b(3) = \{2, 3\}$ , then there are two further subcases:
  - (a) if  $b = (14)(23)\pi$ , and  $4a \neq 2$ , then  $\mathcal{M}(a^2b) = \{2, 4\}$ ,  $\mathcal{M}(a^2ba) = \{3, 4a\}$ ,  $\mathcal{M}(a^2ba^2) = \{1, 3, 4a^2\}$ , and the word  $a^2ba^2$  3-compresses  $\mathcal{A}$ ;
  - (b) if  $b = (145 \dots)(23)\pi$ , then  $\mathcal{M}(a^2b) = \{2, 4\}$ ,  $\mathcal{M}(a^2b^2) = \{3, 5\}$ ,  $\mathcal{M}(a^2b^2a) = \{1, 3, 5a\}$ , and the word  $a^2b^2a$  3-compresses  $\mathcal{A}$ .

3. Let  $Orb_b(3) = \{3, 4\}$ , i.e.,  $b = (34)\pi$ , then in Fig. 17 we draw a P3MSA for this case, proving that either the word  $a^2b^2a$  or  $a^2ba$  or  $aba^2$  or  $ababa$  3-compresses  $\mathcal{A}$ . Observe that if  $1b = 1$ , then  $2b \neq 2$ .

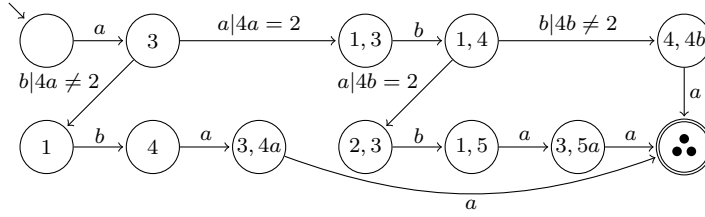


**Fig. 17:** P3MSA for the case  $b = (34)\pi$ .

4. Let  $|Orb_b(3)| \geq 3$ , then there are three main subcases:

(a)  $3b = 1$ .

- i. If  $b = (3124 \dots)\pi$ , then  $\mathcal{M}(a^2b) = \{1, 2\}$ ,  $\mathcal{M}(a^2b^2) = \{2, 4\}$ ,  $\mathcal{M}(a^2b^3) = \{4, 4b\}$ ,  $\mathcal{M}(a^2b^3a) = \{3, 4a, 4ba\}$  and the word  $a^2b^3a$  3-compresses  $\mathcal{A}$ .
- ii. If  $b = (314 \dots)\pi$ , then the P3MSA in Fig. 18 proves that either the word  $a^2bab a^2$  or  $a^2b^2a$  or  $ab^2a^2$  3-compresses  $\mathcal{A}$ . Observe that if  $4a = 2$  and  $4b = 2$ , then  $b$  is of the form  $(31425 \dots)\pi$ , otherwise condition 4. it is not satisfied.



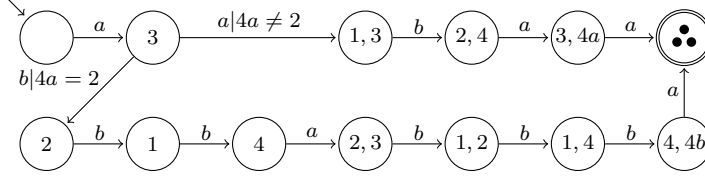
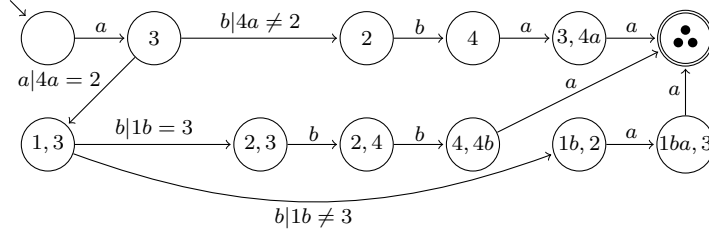
**Fig. 18:** P3MSA for the case  $b = (314 \dots)\pi$ .

(b)  $3b = 2$ .

- i. If  $b = (3214 \dots)\pi$ , then the P3MSA in Fig. 19 proves that either the word  $a^2ba^2$  or  $ab^3ab^3a$  3-compresses  $\mathcal{A}$ .
- ii. If  $b = (324 \dots)\pi$ , then in Fig. 20 we draw a P3MSA for this case, proving that either the word  $ab^2a^2$  or  $a^2b^3a$  or  $a^2ba^2$  3-compresses  $\mathcal{A}$ . Observe that  $1b \neq 3$  implies  $1ba \neq 1$ . Moreover if  $4a = 2$ , then  $1ba \neq 2$ , otherwise the contradiction  $1b = 4$  arises. So when  $4a = 2$  and  $1b \neq 3$  it is  $1ba \notin \{1, 2\}$ .

(c)  $3b = 4$ .

- i. If  $1b \notin \{1, 2\}$ , then  $\mathcal{M}(a^2b) = \{1b, 4\}$ ,  $\mathcal{M}(a^2ba) = \{1ba, 3, 4a\}$ , and the word  $a^2ba$  3-compresses  $\mathcal{A}$ .

Fig. 19: P3MSA for the case  $b = (3214 \dots)\pi$ .Fig. 20: P3MSA for the case  $b = (324 \dots)\pi$ .

- ii. If  $1b \in \{1, 2\}$  and  $4a \neq 2$ , then  $\mathcal{M}(aba) = \{4a, 3\}$  and  $\mathcal{M}(aba^2) = \{4a^2, 1, 3\}$ , and the word  $aba^2$  3-compresses  $\mathcal{A}$ .
- iii. If  $1b \in \{1, 2\}$ ,  $4a \neq 2$  and  $2b \in \{1, 2\}$ , then  $4b \notin \{1, 2, 3\}$ , so  $\mathcal{M}(aba) = \{4ba, 3\}$  (with  $4ba \neq 2$ ) and  $\mathcal{M}(aba^2) = \{4ba^2, 1, 3\}$ , and the word  $aba^2$  3-compresses  $\mathcal{A}$ .
- iv. If  $1b \in \{1, 2\}$ ,  $4a \neq 2$  and  $2b \notin \{1, 2\}$ , then  $\mathcal{M}(aba) = \{2, 3\}$ ,  $\mathcal{M}(abab) = \{2b, 4\}$  and  $\mathcal{M}(ababa) = \{2ba, 2, 3\}$ , and the word  $ababa$  3-compresses  $\mathcal{A}$ .

□

## 4 3-compressible automata without permutations

In this section we characterize proper 3-compressible automata on 2-letter alphabet where no letter acts as permutation.

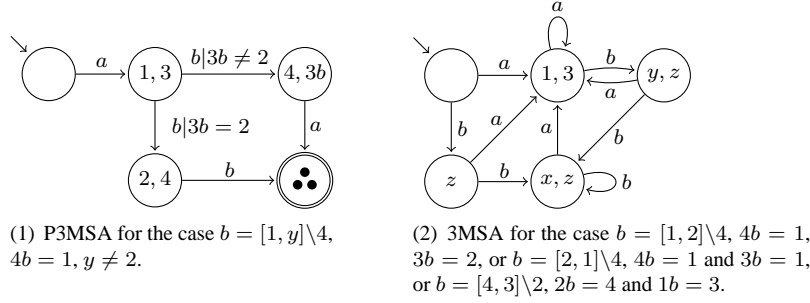
**Proposition 12** *Let  $\mathcal{A}$  a  $(i, j)$ -automaton with  $i \in \{1, 2\}$  and  $j \in \{1, 2, 4\}$ , then  $\mathcal{A}$  is either not 3-compressible or not proper.*

**Proof:** We have to consider five different cases.

1. Let  $\mathcal{A}$  be a  $(1, 1)$ -automaton with  $a = [1, 2, 3] \setminus 1, 2$  and  $b = [x, y, z] \setminus x, y$ :
  - (a) if  $\{1, 2\} \subseteq \{x, y, z\}$  and  $\{x, y\} \subseteq \{1, 2, 3\}$ , then for any  $w \in \{a, b\}^*$ , we have  $\mathcal{M}(wa) = \{1, 2\}$  and  $\mathcal{M}(wb) = \{x, y\}$ , so  $\mathcal{A}$  is not 3-compressible;
  - (b) if  $\{1, 2\} \not\subseteq \{x, y, z\}$ , then  $|\mathcal{M}(ab)| \geq 3$ , so  $\mathcal{A}$  is not proper; similarly if  $\{x, y\} \not\subseteq \{1, 2, 3\}$ , then  $|\mathcal{M}(ba)| \geq 3$ , and again  $\mathcal{A}$  is not proper.

2. Let  $\mathcal{A}$  be a  $(1, 2)$ -automaton with  $a = [1, 2, 3] \setminus 1, 2$  and  $b = [x, y][z, v] \setminus x, z$ :
  - (a) if  $\{1, 2\} \in \{\{x, z\}, \{x, v\}, \{y, z\}, \{y, v\}\}$  and  $\{x, z\} \subseteq \{1, 2, 3\}$ , then for any  $w \in \{a, b\}^*$ , we have  $\mathcal{M}(wa) = \{1, 2\}$  and  $\mathcal{M}(wb) = \{x, z\}$ , so  $\mathcal{A}$  is not 3-compressible;
  - (b) if  $\{1, 2\} \notin \{\{x, z\}, \{x, v\}, \{y, z\}, \{y, v\}\}$ , then  $|\mathcal{M}(ab)| \geq 3$ , so  $\mathcal{A}$  is not proper; similarly if  $\{x, z\} \not\subseteq \{1, 2, 3\}$ , then  $|\mathcal{M}(ba)| \geq 3$ , and again  $\mathcal{A}$  is not proper.
3. Let  $\mathcal{A}$  be a  $(2, 2)$ -automaton with  $a = [1, 2][3, 4] \setminus 1, 3$  and  $b = [x, y][z, v] \setminus x, z$ :
  - (a) if  $\{1, 3\} \in \{\{x, z\}, \{x, v\}, \{y, z\}, \{y, v\}\}$  and  $\{x, z\} \in \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$ , then for all  $w \in \{a, b\}^*$ , we have  $\mathcal{M}(wa) = \{1, 3\}$  and  $\mathcal{M}(wb) = \{x, z\}$ , so  $\mathcal{A}$  is not 3-compressible;
  - (b) if  $\{1, 3\} \notin \{\{x, z\}, \{x, v\}, \{y, z\}, \{y, v\}\}$ , then  $|\mathcal{M}(ab)| \geq 3$ , so  $\mathcal{A}$  is not proper; similarly if  $\{x, z\} \notin \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$ , then  $|\mathcal{M}(ba)| \geq 3$ , and again  $\mathcal{A}$  is not proper.
4. Let  $\mathcal{A}$  be a  $(1, 4)$ -automaton with  $a = [1, 2, 3] \setminus 1, 2$  and  $b = [x, y] \setminus z, zb = x$ :
  - (a) if  $\{x, z\} \not\subseteq \{1, 2, 3\}$ , then  $\mathcal{M}(b^2) = \{x, z\}$  and  $\mathcal{M}(b^2a) = \{1, 2, xa, za\}$ , so  $\mathcal{A}$  is not proper;
  - (b) if  $\{x, z\} = \{1, 2\}$ , then for all  $w \in \{a, b\}^+ \setminus \{b\}$ , we have  $\mathcal{M}(w) = \{1, 2\}$ , so  $\mathcal{A}$  is not 3-compressible;
  - (c) if  $\{x, z\} \in \{\{1, 3\}, \{2, 3\}\}$ , let  $q \in \{1, 2, 3\} \setminus \{x, z\}$ . Then  $\mathcal{M}(ab) = \{z, qb\}$  and  $\mathcal{M}(ab^2) = \{x, z, qb^2\}$ , so  $\mathcal{A}$  is not proper.
5. Let  $\mathcal{A}$  be a  $(2, 4)$ -automaton with  $a = [1, 2][3, 4] \setminus 1, 3$  and  $b = [x, y] \setminus z, zb = x$ :
  - (a) if  $\{x, z\} \notin \{\{1, 3\}, \{1, 4\}, \{2, 3\}, \{2, 4\}\}$ , then  $\mathcal{M}(b^2) = \{x, z\}$  and  $|\mathcal{M}(b^2a)| \geq 3$ , so  $\mathcal{A}$  is not proper;
  - (b) if  $\{x, y\} \cap \{1, 3\} = \emptyset$ , then  $\mathcal{M}(ab) = \{1b, 3b, z\}$ , so  $\mathcal{A}$  is not proper;
  - (c) if  $\{x, z\} = \{1, 3\}$ , then for all  $w \in \{a, b\}^+ \setminus \{b\}$ , we have  $\mathcal{M}(w) = \{1, 3\}$ , so  $\mathcal{A}$  is not 3-compressible;
  - (d) if  $\{x, z\} = \{1, 4\}$  and  $\{x, y\} \cap \{1, 3\} \neq \emptyset$ , we consider two subcases:
    - i. if  $x = 4$  and  $z = 1$ , then  $y$  must be equal to 3. So  $b = [4, 3] \setminus 1$ , then for all  $w \in \{a, b\}^+$ , we have  $\mathcal{M}(wa) = \{1, 3\}$  and  $\mathcal{M}(wb) = \{1, 4\}$ , so  $\mathcal{A}$  is not 3-compressible;
    - ii. if  $x = 1$  and  $z = 4$ , then if  $y \neq 2$  the P3MSA in Fig. 21(1) proves that  $\mathcal{A}$  is not proper. Else, if  $y = 2$  and  $3b \neq 2$ , then the word  $aba$  3-compresses  $\mathcal{A}$  which is not proper. Otherwise, if  $y = 2$  and  $3b = 2$ , then the 3MSA in Fig. 21 proves that  $\mathcal{A}$  is not 3-compressible;
  - (e) if  $\{x, z\} = \{2, 3\}$ , this case reduces to the previous exchanging the state 1 with 3 and 2 with 4;
  - (f) if  $\{x, z\} = \{2, 4\}$  and  $\{x, y\} \cap \{1, 3\} \neq \emptyset$  then  $y \in \{1, 3\}$ . Let  $q \in \{1, 3\} \setminus \{y\}$ .
    - i. If  $b = [2, 1] \setminus 4$  and  $q = 3$ , or  $b = [4, 3] \setminus 2$  and  $q = 1$  then, if  $qb \neq y$  we have  $\mathcal{M}(ab) = \{z, qb\}$  and  $\mathcal{M}(aba) = \{1, 3, qba\}$ , so the automaton is not proper, else if  $qb = y$   $\mathcal{A}$  is not 3-compressible, as shown in Fig. 21(2).





**Fig. 21:** 3MSA for a  $(2, 4)$ -automaton with  $a = [1, 2][3, 4] \setminus 1, 3$  and  $b$  satisfying either condition (d).ii or (f).i.

- ii. Let  $b = [2, 3] \setminus 4$  and  $q = 1$ , or  $b = [4, 1] \setminus 2$  and  $q = 3$  then either the word  $ab^2$  or  $aba$  3-compresses  $\mathcal{A}$ , so it is not proper.

□

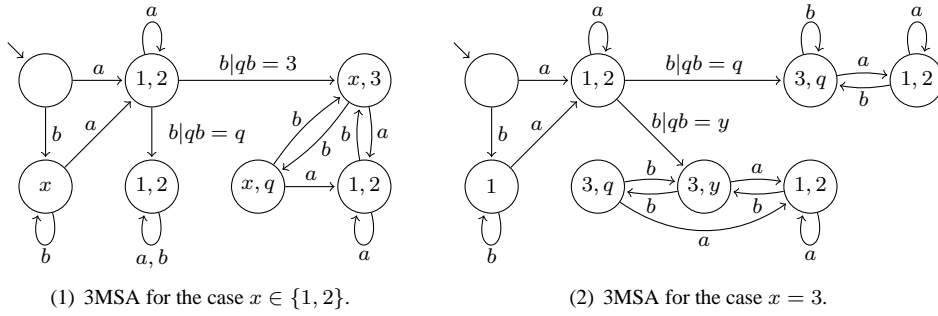
**Proposition 13** Let  $\mathcal{A}$  be a  $(1, 3)$ -automaton with  $a = [1, 2, 3] \setminus 1, 2$  and  $b = [x, y] \setminus x$ . Then  $\mathcal{A}$  is 3-compressible and proper if, and only if, the following conditions hold:

1.  $x \in \{1, 2, 3\}$ , and
2.  $\{1, 2\} \cap \{x, y\} \neq \emptyset$ , and
3. for all  $q \in \{1, 2\} \setminus \{x\}$ ,  $qb \in \{1, 2, 3\}$ , and  $Orb_b(q) \not\subseteq \{1, 2, 3\}$ .

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then the word  $ab^2a$  3-compresses  $\mathcal{A}$ .

**Proof:** Let  $\mathcal{A}$  be a  $(1, 3)$ -automaton that does not satisfy one of the conditions 1. – 3.

1. Let  $x \notin \{1, 2, 3\}$ , then the word  $ba$  3-compresses  $\mathcal{A}$ , and so it is not proper.
2. Let  $\{1, 2\} \cap \{x, y\} = \emptyset$ , then the word  $ab$  3-compresses  $\mathcal{A}$ , and so it is not proper.
3. Let for all  $q \in \{1, 2\} \setminus \{x\}$  be  $qb \notin \{1, 2, 3\}$  or  $Orb_b(q) \subseteq \{1, 2, 3\}$ , and  $x \in \{1, 2, 3\}$ , and  $\{1, 2\} \cap \{x, y\} \neq \emptyset$ .
  - (a) Let  $qb \notin \{1, 2, 3\}$ , then if  $x \in \{1, 2\}$ , we have  $\mathcal{M}(ab) = \{qb, x\}$  and  $\mathcal{M}(aba) = \{1, 2, qba\}$ , else if  $x = 3$ , we choose  $q \in \{1, 2\} \setminus \{y\}$ , and again  $\mathcal{M}(ab) = \{qb, x\}$  and  $\mathcal{M}(aba) = \{1, 2, qba\}$ . So if  $qb \notin \{1, 2, 3\}$  the word  $aba$  3-compresses  $\mathcal{A}$  which is not proper.
  - (b) So let  $qb \in \{1, 2, 3\}$  and  $Orb_b(q) \subseteq \{1, 2, 3\}$ .
    - i. If  $x \in \{1, 2\}$ , then  $qb \in \{q, 3\}$  and  $3b = q$ , so  $\mathcal{A}$  is not 3-compressible, as shown in Fig. 22(1).
    - ii. If  $x = 3$ , then  $y \in \{1, 2\}$  and  $\{1b, 2b\} = \{1, 2\}$ , and  $\mathcal{A}$  is not 3-compressible, as shown in Fig. 22(2), where in this case  $q \in \{1, 2\} \setminus \{y\}$ .



**Fig. 22:** 3MSA for automata that do not satisfy condition 3. of Proposition 13 with  $zb \in \{1, 2, 3\}$  and  $Orb_b(z) \subseteq \{1, 2, 3\}$ .

Conversely, suppose  $x \in \{1, 2, 3\}$ ,  $\{1, 2\} \cap \{x, y\} \neq \emptyset$ , and for all  $q \in \{1, 2\} \setminus \{x\}$ ,  $qb \in \{1, 2, 3\}$  and  $Orb_b(q) \not\subseteq \{1, 2, 3\}$ . If  $x \in \{1, 2\}$ , let  $q \in \{1, 2\} \setminus \{x\}$ , else let  $q \in \{1, 2\} \setminus \{y\}$ ; then  $\mathcal{M}(ab) = \{x, qb\}$ ,  $\mathcal{M}(ab^2) = \{x, qb^2\}$ , and, as  $qb^2 \notin \{1, 2, 3\}$ ,  $\mathcal{M}(ab^2a) = \{1, 2, qb^2a\}$ , and the word  $ab^2a$  3-compresses  $\mathcal{A}$ .  $\square$

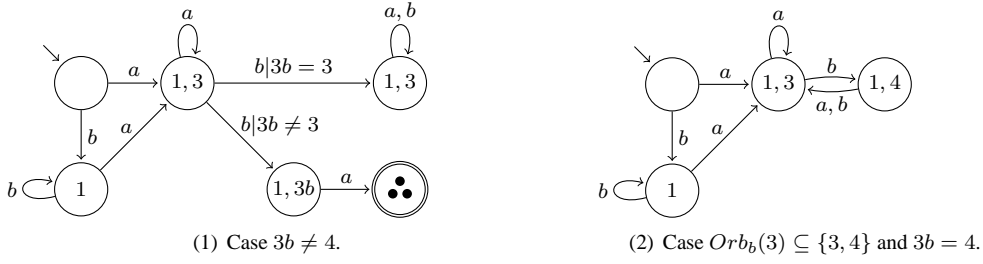
**Proposition 14** Let  $\mathcal{A}$  be a  $(2, 3)$ -automaton with  $a = [1, 2][3, 4] \setminus 1, 3$  and  $b = [x, y] \setminus x$ . Then  $\mathcal{A}$  is 3-compressible and proper if, and only if, the following conditions hold:

1.  $x \in \{1, 2, 3, 4\}$ , and
2. if  $x = 1$ , then  $3b = 4$  and  $Orb_b(3) \not\subseteq \{3, 4\}$ , and
3. if  $x = 2$ , then  $y \in \{1, 3\}$  and
  - (a) if  $y = 1$ , then  $3b = 4$  and  $4b \neq 3$ , and
  - (b) if  $y = 3$ , then  $1b \in \{3, 4\}$ , and
4. if  $x = 3$ , then  $1b = 2$  and  $Orb_b(1) \not\subseteq \{1, 2\}$ , and
5. if  $x = 4$ , then  $y \in \{1, 3\}$  and
  - (a) if  $y = 1$ , then  $3b \in \{1, 2\}$ , and
  - (b) if  $y = 3$ , then  $1b = 2$  and  $2b \neq 1$ .

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then either the word  $ab^2a$  or  $ab^3a$  3-compresses  $\mathcal{A}$ .

**Proof:** Let  $\mathcal{A}$  be a  $(2, 3)$ -automaton that does not satisfy one of the conditions 1. – 5.

1. Let  $x \notin \{1, 2, 3, 4\}$ , then the word  $ba$  3-compresses  $\mathcal{A}$ , and so it is not proper.
2. Let  $x = 1$ ,
  - (a) if  $3b \neq 4$ , then if  $3b = 3$   $\mathcal{A}$  is not 3-compressible, while if  $3b \neq 3$  the word  $aba$  3-compresses  $\mathcal{A}$  that it is not proper, as shown in Fig. 23(1);

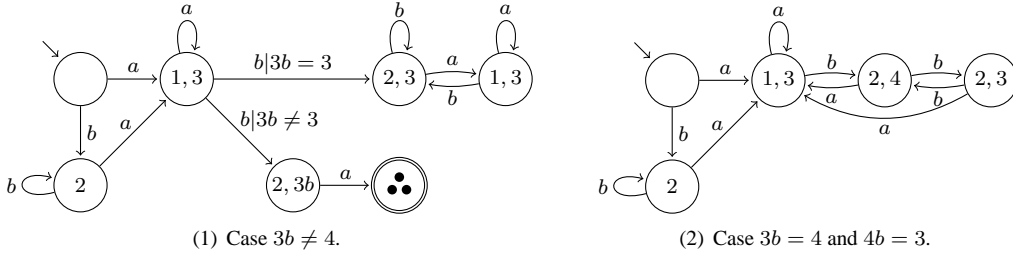


**Fig. 23:** 3MSA for automata that do not satisfy condition 2. of Proposition 14.

(b) if  $3b = 4$  and  $Orb_b(3) \subseteq \{3, 4\}$ , then  $\mathcal{A}$  is not 3-compressible, as shown in Fig. 23(2).

3. Let  $x = 2$ ,

- (a) if  $y \notin \{1, 3\}$ , then the word  $ab$  3-compresses  $\mathcal{A}$ , and so it is not proper;
- (b) if  $y = 1$ , then if  $3b \neq 4$ , then  $\mathcal{A}$  is not 3-compressible, as shown in Fig. 24(1); while if  $3b = 4$  and  $4b = 3$ , then  $\mathcal{A}$  is not 3-compressible, as shown in Fig. 24(2);



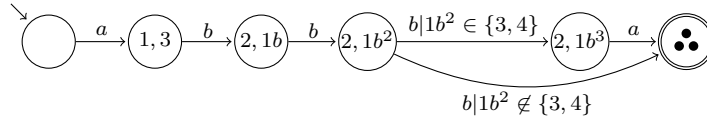
**Fig. 24:** 3MSA for automata that do not satisfy condition 3. of Proposition 14 with  $y = 1$ .

- (c) if  $y = 3$  and  $1b \notin \{3, 4\}$ , then  $\mathcal{M}(ab) = \{2, 1b\}$ , and  $\mathcal{M}(aba) = \{1, 3, 1ba\}$ , so the word  $aba$  3-compresses  $\mathcal{A}$  that it is not proper.

The cases  $x = 3$  and  $x = 4$  reduce to case 2. and 3., respectively.

Conversely, suppose  $x \in \{1, 2, 3, 4\}$ .

1. Let  $x = 1$ ,  $3b = 4$  and  $Orb_b(3) \not\subseteq \{3, 4\}$ , this implies  $4b = 3b^2 \notin \{3, 4\}$ . Then  $\mathcal{M}(ab) = \{1, 4\}$ ,  $\mathcal{M}(ab^2) = \{1, 4b\}$  and  $\mathcal{M}(ab^2a) = \{1, 3, 4ba\}$ , so  $ab^2a$  3-compresses  $\mathcal{A}$ .
2. Let  $x = 2$ .
  - (a) If  $y = 1$ ,  $3b = 4$  and  $4b \neq 3$ , then  $\mathcal{M}(ab) = \{2, 4\}$ ,  $\mathcal{M}(ab^2) = \{2, 4b\}$  and  $\mathcal{M}(ab^2a) = \{1, 3, 4ba\}$ , so  $ab^2a$  3-compresses  $\mathcal{A}$ .
  - (b) If  $y = 3$  and  $1b \in \{3, 4\}$ , then either the word  $ab^2a$  or  $ab^3a$  3-compresses  $\mathcal{A}$ , as shown in Fig. 25.



**Fig. 25:** P3MSA for the case  $b = [2, 3] \setminus 2$  and  $1b \in \{3, 4\}$ .

The cases  $x = 3$  and  $x = 4$  reduce to cases 1. and 2., respectively.

□

The following lemma is straightforward.

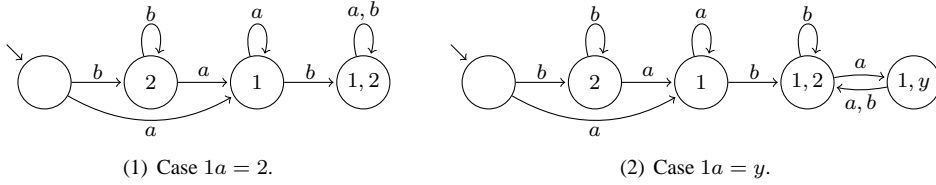
**Lemma 15** *Let  $\mathcal{A}$  be a  $(3, 3)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus x$ . Then  $x \neq 1$  and  $\{x, y\} \neq \{1, 2\}$ .*

**Lemma 16** *Let  $\mathcal{A}$  be a  $(3, 3)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [2, y] \setminus 2$  with  $y \neq 1$ . Then:*

1.  $\{1ba, 1b^2a, 1b^3a\} \not\subseteq \{2, y\}$ , or
2.  $Orb_a(1ba) \not\subseteq \{2, y\}$ .

**Proof:** Let  $\mathcal{A}$  be a  $(3, 3)$ -automaton that does not satisfy the above conditions, we prove that it is not 3-compressible or not proper.

1. Let  $1b = 1$  and  $Orb_a(1ba) \subseteq \{2, y\}$ .
  - (a) If  $1ba = 1a = 2a = 2$ , then the 3MSA in Fig. 26(1) proves that  $\mathcal{A}$  is not 3-compressible;
  - (b) If  $1ba = 1a = 2a = y$ , then  $ya = 1ba^2 = 2$  and the 3MSA in Fig. 26(2) proves that  $\mathcal{A}$  is not 3-compressible.



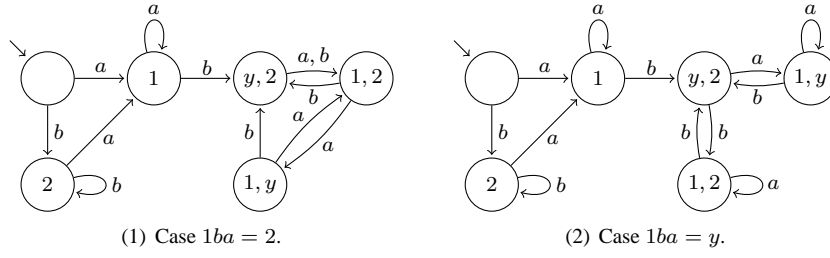
**Fig. 26:** 3MSA for automata that do not satisfy conditions 1. and 2. of Lemma 16 with  $1b = 1$ .

2. Let  $1b \neq 1$ ,  $\{1ba, 1b^2a, 1b^3a\} \subseteq \{2, y\}$  and  $Orb_a(1ba) \subseteq \{2, y\}$ . If  $1ba = 1b^2a$  or  $1b^2a = 1b^3a$ , then for all  $h > 0$   $1b^h \neq 2$ , then  $1 = 1b$ , against the hypothesis. If  $1b^2a = 1b^3a$ , as  $1b^2 \neq 2$ , then  $1b^2 = 1b^3$ , and then  $1 = 1b$ , against the hypothesis. So, as  $|\{1ba, 1b^2a, 1b^3a\}| \leq 2$ ,  $1ba = 1b^3a$ , then  $1b = 1b^3$  and  $1 = 1b^2$ .
  - (a) If  $1ba = 2$ , then  $1b^2a = y$ , hence  $1a = 2a = y$ ,  $ya = 2$  and  $1b = y$ . The 3MSA in Fig. 27(1) proves that  $\mathcal{A}$  is not 3-compressible;
  - (b) If  $1ba = y$ , then  $1b^2a = 1a = 2a = 2$ . Moreover,  $1ba^2 = y$ , otherwise  $1ba^2 = 2$  gives the contradiction  $1ba = 1$ , hence  $ya = y$ ,  $1b = y$  and  $yb = 1b^2 = 1$ . The 3MSA in Fig. 27(2) proves that  $\mathcal{A}$  is not 3-compressible.

□

**Corollary 17** *Let  $\mathcal{A}$  be a  $(3, 3)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, 1] \setminus x$  with  $x \neq 2$ . Then:*

1.  $\{xab, xa^2b, xa^3b\} \not\subseteq \{1, 2\}$ , or



**Fig. 27:** 3MSA for automata that do not satisfy conditions 1. and 2. of Lemma 16 with  $1b \neq 1$ .

2.  $Orb_b(xab) \not\subseteq \{1, 2\}$ .

**Proof:** It is a straightforward consequence of the previous lemma, simply replacing  $a$  with  $b$ , 1 with 2, 2 with  $y$  and  $x$  with 1.  $\square$

**Lemma 18** Let  $\mathcal{A}$  be a proper  $(3, 3)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus x$  with  $x \notin \{1, 2\}$  and  $y \neq 1$ . Then all the following conditions hold:

1.  $1b \in \{1, 2\}$ , and
2.  $xa \in \{x, y\}$ , and
3.  $Orb_b(1) \not\subseteq \{1, 2\}$  or  $Orb_a(x) \not\subseteq \{x, y\}$ .

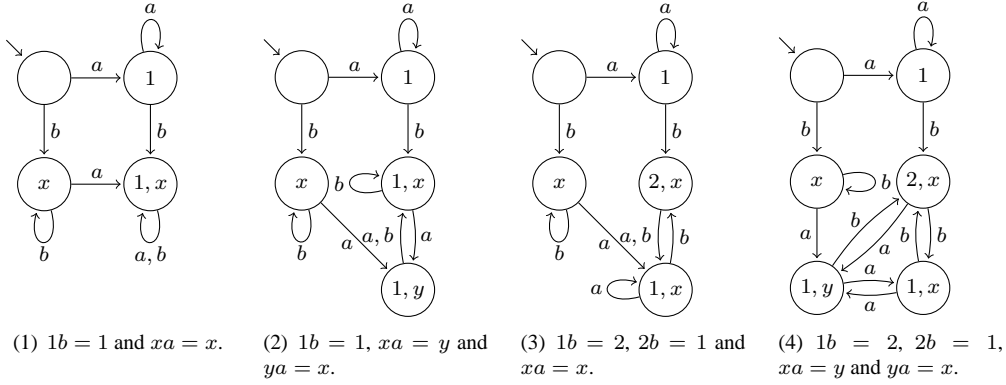
**Proof:** Let  $\mathcal{A}$  be a  $(3, 3)$ -automaton that does not satisfy the above conditions, we prove that it is not 3-compressible or not proper.

1. If  $1b \notin \{1, 2\}$ , then  $\mathcal{A}$  is not proper, in fact  $\mathcal{M}(ab) = \{x, 1b\}$  and  $|\mathcal{M}(aba)| = 3$ ; similarly
2. if  $1b \in \{1, 2\}$  and  $xa \notin \{x, y\}$ , then  $\mathcal{A}$  is not proper, in fact  $\mathcal{M}(ba) = \{1, xa\}$  and  $|\mathcal{M}(bab)| = 3$ .
3. If  $1b \in \{1, 2\}$ ,  $xa \in \{x, y\}$ ,  $Orb_b(1) \subseteq \{1, 2\}$  and  $Orb_a(x) \subseteq \{x, y\}$ , then
  - (a) if  $1b = 1$  and  $xa = x$ , the 3MSA of  $\mathcal{A}$  is in Fig. 28(1);
  - (b)  $1b = 1$  and  $xa = y$  (and then  $ya = x$ ), the 3MSA of  $\mathcal{A}$  is in Fig. 28(2);
  - (c)  $1b = 2$  (and then  $2b = 1$ ) and  $xa = x$ , the 3MSA of  $\mathcal{A}$  is in Fig. 28(3);
  - (d)  $1b = 2$  (and then  $2b = 1$ ) and  $xa = y$  (and then  $ya = x$ ) the 3MSA of  $\mathcal{A}$  is in Fig. 28(4).

In all the subcases,  $\mathcal{A}$  is not 3-compressible.

$\square$

**Proposition 19** Let  $\mathcal{A}$  be a  $(3, 3)$ -automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus x$ . Then  $\mathcal{A}$  is proper and 3-compressible if, and only if, one of the following conditions holds:



**Fig. 28:** 3MSA for automata that do not satisfy conditions 3. of Lemma 18

1.  $b = [2, y] \setminus 2$ ,  $y \neq 1$ , and either  $\{1ba, 1b^2a, 1b^3a\} \not\subseteq \{2, y\}$  or  $\text{Orb}_a(1ba) \not\subseteq \{2, y\}$ ;
2.  $b = [x, 1] \setminus x$ ,  $x \neq 2$ , and either  $\{xba, xb^2a, xb^3a\} \not\subseteq \{1, 2\}$  or  $\text{Orb}_b(xab) \not\subseteq \{1, 2\}$ ;
3.  $b = [x, y] \setminus x$ ,  $x \notin \{1, 2\}$ ,  $y \neq 1$ ,  $1b \in \{1, 2\}$ ,  $xa \in \{x, y\}$ , and either  $\text{Orb}_b(1) \not\subseteq \{1, 2\}$  or  $\text{Orb}_a(x) \not\subseteq \{x, y\}$ .

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then one of the words  $abab$  or  $ab^2ab$  or  $ab^3ab$  or  $aba^2b$  or  $aba^3b$  or  $baba$  or  $ba^2ba$  or  $ba^3ba$  or  $bab^2a$  or  $bab^3a$  3-compresses  $\mathcal{A}$ .

**Proof:** If  $\mathcal{A}$  is a 3-compressible (3, 3)-automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus x$ , then, from the previous lemmata and corollary, one of the above conditions must hold.

Conversely, we find for any automaton satisfying conditions 1. – 3. a (short) 3-compressing word.

1. Let  $b = [2, y] \setminus 2$ ,  $y \neq 1$ ,
  - (a) if  $\{1ba, 1b^2a, 1b^3a\} \not\subseteq \{2, y\}$ , then either the word  $abab$  or  $ab^2ab$  or  $ab^3ab$  3-compresses  $\mathcal{A}$ , since  $\mathcal{M}(aba) = \{1, 1ba\}$ ,  $\mathcal{M}(ab^2a) = \{1, 1b^2a\}$  and  $\mathcal{M}(ab^3a) = \{1, 1b^3a\}$ ;
  - (b) if  $\text{Orb}_a(1ba) \not\subseteq \{2, y\}$ , then  $\{1ba, 1ba^2, 1ba^3\} \not\subseteq \{2, y\}$  and either the word  $abab$  or  $aba^2b$  or  $aba^3b$  3-compresses  $\mathcal{A}$ , since  $\mathcal{M}(aba) = \{1, 1ba\}$ ,  $\mathcal{M}(aba^2) = \{1, 1ba^2\}$  and  $\mathcal{M}(aba^3) = \{1, 1ba^3\}$ .
2. Let  $b = [x, 1] \setminus x$ ,  $x \neq 2$ , and either  $\{xba, xb^2a, xb^3a\} \not\subseteq \{1, 2\}$  or  $\text{Orb}_b(xab) \not\subseteq \{1, 2\}$ . This case reduces to the previous one replacing  $a$  with  $b$ , 1 with 2, 2 with  $y$  and  $x$  with 1, then either the word  $baba$  or  $ba^2ba$  or  $ba^3ba$  or  $bab^2a$  or  $bab^3a$  3-compresses  $\mathcal{A}$ .
3. Let  $b = [x, y] \setminus x$ ,  $x \notin \{1, 2\}$ ,  $y \neq 1$ ,  $1b \in \{1, 2\}$ ,  $xa \in \{x, y\}$ . If  $\text{Orb}_b(1) \not\subseteq \{1, 2\}$ , then  $1b = 2$  and  $2b \notin \{1, 2\}$ , and so the word  $ab^2a$  3-compresses  $\mathcal{A}$ , as  $\mathcal{M}(ab^2) = \{x, 2b\}$ ; similarly, if  $\text{Orb}_a(x) \not\subseteq \{x, y\}$ , exchanging  $b$  and  $a$  and so  $x$  with 1 and  $y$  with 2, the word  $ba^2b$  3-compresses  $\mathcal{A}$ .

□

**Lemma 20** *Let  $\mathcal{A}$  be a  $(3, 4)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . If  $\mathcal{A}$  is proper then all the following conditions hold:*

1.  $\{1, 2\} \cap \{x, z\} \neq \emptyset$ ,
2.  $z \in \{1, 2\}$  or  $\{1, za\} \cap \{x, y\} \neq \emptyset$ ,
3.  $\{1, 1b\} \cap \{x, y\} \neq \emptyset$ ,
4.  $1 \in \{x, y\}$  or  $\{z, 1b\} \cap \{1, 2\} \neq \emptyset$ .

**Proof:** If  $\{1, 2\} \cap \{x, z\} = \emptyset$ , then  $|\mathcal{M}(b^2a)| = 3$ ; if  $z \notin \{1, 2\}$  and  $\{1, za\} \cap \{x, y\} = \emptyset$ , then  $|\mathcal{M}(bab)| = 3$ ; if  $\{1, 1b\} \cap \{x, y\} = \emptyset$ , then  $|\mathcal{M}(ab^2)| = 3$ ; if  $1 \notin \{x, y\}$  and  $\{z, 1b\} \cap \{1, 2\} = \emptyset$ , then  $|\mathcal{M}(aba)| = 3$ . So each automaton that does not satisfy one of the conditions of the lemma is 3-compressed by a word of length 3, and then it is not proper. □

**Corollary 21** *Let  $\mathcal{A}$  be a proper  $(3, 4)$  3-compressible automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . The following conditions hold:*

1.  $b = [x, y] \setminus 1$ , or
2.  $b = [x, y] \setminus 2$  and  $\{1, 1b\} \cap \{x, y\} \neq \emptyset$ , or
3.  $b = [1, y] \setminus z$  or  $b = [2, 1] \setminus z$  and in both cases  $z \neq 2$ .

**Proof:** It is a straightforward consequence of the previous lemma. □

**Proposition 22** *Let  $\mathcal{A}$  be a  $(3, 4)$ -automaton with  $a = [1, 2] \setminus 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . Then  $\mathcal{A}$  is proper and 3-compressible if, and only if, the following conditions hold:*

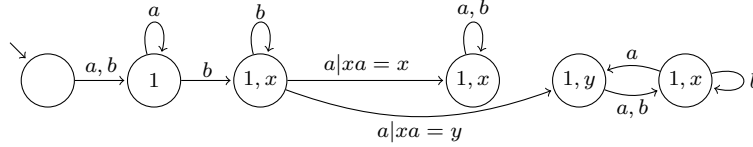
1. if  $b = [x, y] \setminus 1$ , then  $\text{Orb}_a(x) \not\subseteq \{x, y\}$ ,
2. if  $z = 2$  and
  - (a)  $b = [1, y] \setminus 2$ , then  $\text{Orb}_a(2) \not\subseteq \{2, y\}$  or  $2ab \notin \{1, y\}$ ,
  - (b)  $b = [x, 1] \setminus 2$ , then  $\{2a, xa\} \neq \{2, x\}$  or  $\{2ab, xab\} \neq \{1, x\}$
  - (c)  $b = [x, y] \setminus 2$ ,  $1b = y$ , and  $1 \notin \{x, y\}$ , then  $\{xa, ya\} \neq \{x, y\}$ ,
3. if  $z \notin \{1, 2\}$  and
  - (a)  $b = [1, y] \setminus z$  and  $y \neq 2$ , then  $za \neq z$ ,
  - (b)  $b = [x, y] \setminus z$  and  $\{x, y\} = \{1, 2\}$ , then  $|\text{Orb}_a(z)| > 2$  or  $zab \notin \{1, 2\}$ .

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then either the word  $b^2ab^2$  or  $b^2a^2b^2$  or  $b^2a^3b^2$  or  $b^2abab^2$  3-compresses  $\mathcal{A}$ .



**Proof:** First observe that from the previous corollary a proper 3-compressible  $(3, 4)$ -automaton always satisfies the antecedent of one of the above conditions, so all the possible cases are taken into account. We start proving that a  $(3, 4)$ -automaton that does not satisfy the conditions 1. – 3. is not 3-compressible.

1. Let condition 1. be false, i.e.,  $b = [x, y] \setminus 1$  and  $Orb_a(x) \subseteq \{x, y\}$ . The 3MSA in Fig. 29 proves that  $\mathcal{A}$  is not 3-compressible.

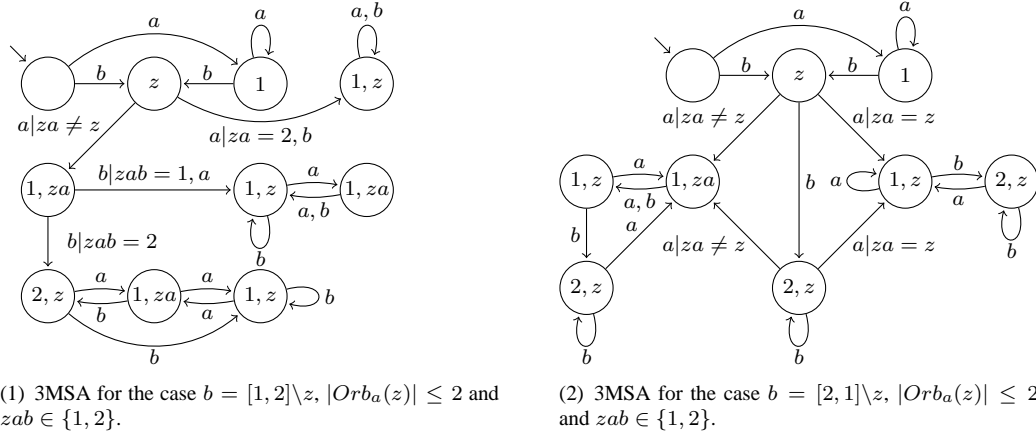


**Fig. 29:** 3MSA for the case in which condition 1. of Proposition 22 is false.

2. Let  $z = 2$ .

- (a) Let condition 2.(a) be false, i.e.,  $b = [1, y] \setminus 2$ ,  $Orb_a(2) \subseteq \{2, y\}$  and  $2ab \in \{1, y\}$ . Observe that if  $2a \neq 2$ , then  $2a = y$  and  $ya = 2$ . Since  $y \neq 2$ , then  $yb = 2ab \neq 1$ , and so  $2ab = y$ . The 3MSA in Fig. 30(1) proves that  $\mathcal{A}$  is not 3-compressible.
- (b) Let condition 2.(b) be false, i.e.,  $b = [x, 1] \setminus 2$ ,  $\{2a, xa\} = \{2, x\}$  and  $\{2ab, xab\} = \{1, x\}$ . Observe that if  $2a = 2$ , then  $2ab = 2b = x$  and so  $xab = 1$ . If  $xa = 2$ , then  $xab = 2b = x$  and so  $2ab = 1$ . The 3MSA in Fig. 30(2) proves that  $\mathcal{A}$  is not 3-compressible.
- (c) Let condition 2.(c) be false, i.e.,  $b = [x, y] \setminus 2$ ,  $1b = y$ ,  $1 \notin \{x, y\}$  and  $\{xa, ya\} = \{x, y\}$ . The 3MSA in Fig. 30(3) proves that  $\mathcal{A}$  is not 3-compressible.

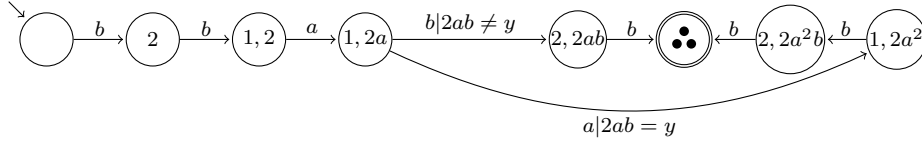




**Fig. 31:** 3MSA for the case in which condition 3.(b) of Proposition 22 is false.

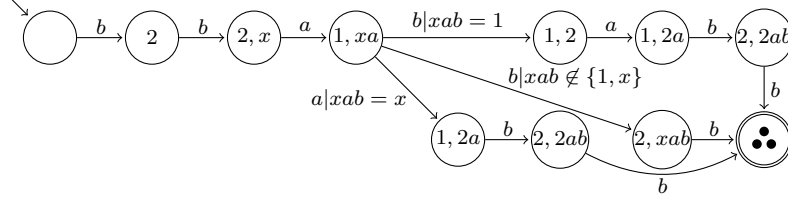
Conversely, we find for any automaton satisfying conditions 1. – 3. a (short) 3-compressing word.

1. Let  $b = [x, y] \setminus 1$  and  $Orb_a(x) \not\subseteq \{x, y\}$ , whence  $xa \neq x$ . Then either  $xa$  or  $xa^2$  are different from  $y$ , and so either the word  $b^2ab$  or  $b^2a^2b$  3-compresses  $\mathcal{A}$ .
2. Let  $z = 2$ , we consider various subcases.
  - (a) Let  $b = [1, y] \setminus 2$ . Then  $2a \neq 2$ , otherwise  $Orb_a(2) = \{2\}$  and  $2ab = 1$ , against the hypothesis. Hence  $2ab = y$  and  $2a^2b \neq y$ . The P3MSA in Fig. 32 proves that either the word  $b^2ab^2$  or  $b^2a^2b^2$  3-compresses  $\mathcal{A}$ ;



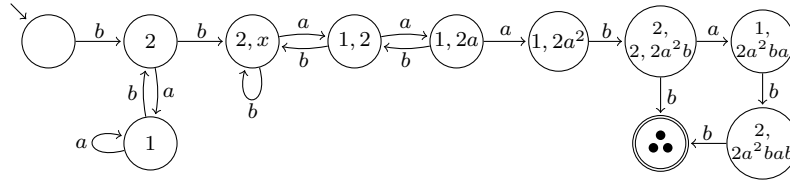
**Fig. 32:** P3MSA for the case 2.(a) of Proposition 22.

- (b) Let  $b = [x, 1] \setminus 2$ :
  - i. if  $\{2ab, xab\} \neq \{1, x\}$ , then the P3MSA in Fig. 33 proves that the word  $b^2ab^2$  or  $b^2a^2b^2$  or  $b^2abab^2$  3-compresses  $\mathcal{A}$ . Observe that if  $xab = x$ , then  $xa = 2$  and  $2ab \notin \{1, x\}$ .
  - ii. if  $2ab = x$ ,  $xab = 1$ , and  $\{2a, xa\} \neq \{2, x\}$ , then  $2ab = 2b$ ,  $2a = 2$  and then  $xa \notin \{2, x\}$ . So  $\mathcal{M}(b^2a) = \{1, xa\}$ ,  $\mathcal{M}(b^2a^2) = \{1, xa^2\}$  and  $\mathcal{M}(b^2a^2b) = \{2, xa^2b\}$ . If  $xa^2b = x = 2b$ , then  $xa^2 = 2a$ , and so  $xa = 2$ , against the hypothesis. Else, if  $xa^2b = 1 = xab$ , then  $xa^2 = xa$ , and so  $xa = x$ , against the hypothesis. Then it follows that  $|\mathcal{M}(b^2a^2b^2)| = 3$  and the word  $b^2a^2b^2$  3-compresses  $\mathcal{A}$ .



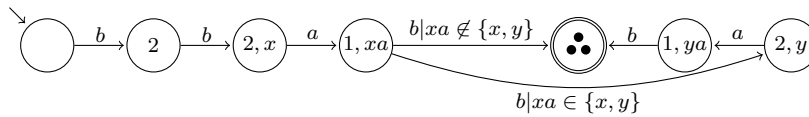
**Fig. 33:** P3MSA for the case  $b = [x, 1] \setminus 2$  and  $\{2ab, xab\} \neq \{1, x\}$ .

- iii. if  $2ab = 1$ ,  $xab = x$ , and  $\{2a, xa\} \neq \{2, x\}$ , then  $xab = 2b$ ,  $xa = 2$  and then  $2a \notin \{2, x\}$ . Then the P3MSA in Fig. 34 proves that any word belonging to the language  $\mathcal{L} = b(a^+b)^*(ba)^+a(ba)^*ab(b+ab^2)$  (and in particular  $b^2a^3b^2$ ) 3-compresses  $\mathcal{A}$ . Indeed, if  $2a^2b = x = 2b$ , then  $2a^2 = 2$ , against the hypothesis, else, if  $2a^2b = 1 = 2ab$ , then  $2a^2 = 2a$ , and so  $2a = 2$ , against the hypothesis, and so  $2a^2b \notin \{1, x\}$ . Moreover, if  $2a^2bab = 1 = 2ab$ , then  $2a^2ba = 2a$ , and so  $2a^2b = 2$ , against the hypothesis else, if  $2a^2bab = x = xab$ , then  $xa^2ba = xa$ , and so  $2a^2b = x$ , against the hypothesis, and so  $2a^2bab \notin \{1, x\}$ .



**Fig. 34:** P3MSA for the case  $b = [x, 1] \setminus 2$ ,  $2ab = 1$ ,  $xab = x$  and  $\{2a, xa\} \neq \{2, x\}$ .

- (c) Let  $b = [x, y] \setminus 2$ ,  $1b = y$ ,  $1 \notin \{x, y\}$  and  $\{xa, ya\} \neq \{x, y\}$ . Observe that if  $xa \in \{x, y\}$ , then  $ya \notin \{x, y\}$ . The P3MSA in Fig. 35 proves that in this case either the word  $b^2ab$  or  $b^2abab$  3-compresses  $\mathcal{A}$ .

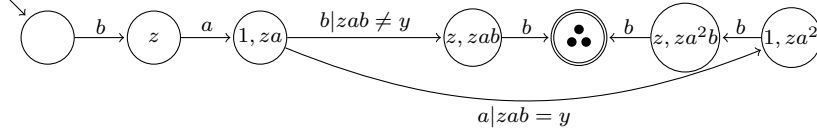


**Fig. 35:** P3MSA for the case  $b = [x, y] \setminus 2$ ,  $1b = y$ ,  $1 \notin \{x, y\}$  and  $\{xa, ya\} \neq \{x, y\}$ .

3. Let  $z \notin \{1, 2\}$ , we consider two main subcases.

- (a) Let  $b = [1, y] \setminus z$ ,  $y \neq 2$  and  $za \neq z$ . Then  $\mathcal{M}(ba) = \{1, za\}$ ,  $\mathcal{M}(bab) = \{z, zab\}$ . If  $zab \neq y$ , then  $|\mathcal{M}(bab^2)| = 3$ , else if  $zab = y$  then  $|\mathcal{M}(baba)| = 3$ , hence either the word  $bab^2$  or  $baba$  3-compresses  $\mathcal{A}$ .
- (b) Let  $\{x, y\} = \{1, 2\}$ . We consider two subcases.

- i. Let  $b = [x, y] \setminus z$  and  $|Orb_a(z)| > 2$ . In particular  $za \neq z$ ,  $zab \neq za^2b$  and  $z \neq za^2$ . If  $zab = y$ , then  $za^2b \notin \{zb, zab\} = \{1, 2\}$ . The P3MSA in Fig. 36 proves that either the word  $ba^2b^2$  or  $bab^2$  3-compresses  $\mathcal{A}$ .



**Fig. 36:** P3MSA for the case  $b = [x, y] \setminus z$ ,  $\{x, y\} = \{1, 2\}$  and  $|Orb_a(z)| > 2$ .

- ii. Let  $b = [x, y] \setminus z$  and  $zab \notin \{1, 2\}$ . Then  $\mathcal{M}(bab) = \{z, zab\}$  and  $|\mathcal{M}(bab^2)| = |\mathcal{M}(baba)| = 3$ , both the words  $bab^2$  and  $baba$  3-compress  $\mathcal{A}$ .

□

**Lemma 23** Let  $\mathcal{A}$  be a  $(4, 4)$  3-compressible automaton with  $a = [1, 2] \setminus 3$ ,  $3a = 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . If  $\mathcal{A}$  is proper then all the following conditions hold:

1.  $\{1, 2\} \cap \{x, z\} \neq \emptyset$ ,
2.  $z \in \{1, 2\}$  or  $\{3, za\} \cap \{x, y\} \neq \emptyset$ ,
3.  $\{3, 3b\} \cap \{x, y\} \neq \emptyset$ ,
4.  $\{1, 3\} \cap \{x, y\} \neq \emptyset$ ,
5.  $3 \in \{x, y\}$  or  $\{z, 3b\} \cap \{1, 2\} \neq \emptyset$ ,
6.  $\{z, za\} \cap \{1, 2\} \neq \emptyset$ .

**Proof:** If  $\{1, 2\} \cap \{x, z\} = \emptyset$ , then  $\mathcal{M}(b^2) = \{x, z\}$  and  $|\mathcal{M}(b^2a)| = 3$ ; if  $z \notin \{1, 2\}$  and  $\{3, za\} \cap \{x, y\} = \emptyset$ , then  $\mathcal{M}(ba) = \{3, za\}$  and  $|\mathcal{M}(bab)| = 3$ ; if  $\{3, 3b\} \cap \{x, y\} = \emptyset$ , then  $\mathcal{M}(ab) = \{3, zb\}$  and  $|\mathcal{M}(ab^2)| = 3$ ; if  $\{1, 3\} \cap \{x, y\} = \emptyset$ , then  $\mathcal{M}(a^2) = \{1, 3\}$  and  $|\mathcal{M}(a^2b)| = 3$ ; if  $3 \notin \{x, y\}$  and  $\{z, 3b\} \cap \{1, 2\} = \emptyset$ , then  $\mathcal{M}(ab) = \{z, 3b\}$  and  $|\mathcal{M}(a^2b)| = 3$ ; if  $\{z, za\} \cap \{1, 2\} = \emptyset$ , then  $\mathcal{M}(ba) = \{z, 3b\}$  and  $|\mathcal{M}(bab)| = 3$ . Then each automaton that does not satisfies one of the above is not proper. □

**Corollary 24** Let  $\mathcal{A}$  be a proper  $(4, 4)$  3-compressible automaton with  $a = [1, 2] \setminus 3$ ,  $3a = 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . The following conditions hold:

1. if  $z = 1$ , then  $3 \in \{x, y\}$ ,
2. if  $z = 2$ , then  $3 \in \{x, y\}$  or it is  $1 \in \{x, y\}$  and  $3b \in \{x, y\}$ ,
3. if  $z \notin \{1, 2\}$ , then  $x \in \{1, 2\}$  and  $za \in \{1, 2\}$ .

**Proof:** It is a straightforward consequence of the previous lemma. Observe that not all the conditions of Lemma 23 are applied, so some automaton satisfying the conditions of the corollary could possibly be not proper or not 3-compressible.  $\square$

**Proposition 25** *Let  $\mathcal{A}$  be a  $(4, 4)$ -automaton with  $a = [1, 2] \setminus 3$ ,  $3a = 1$  and  $b = [x, y] \setminus z$ ,  $zb = x$ . Then  $\mathcal{A}$  is proper and 3-compressible if, and only if, the following conditions hold:*

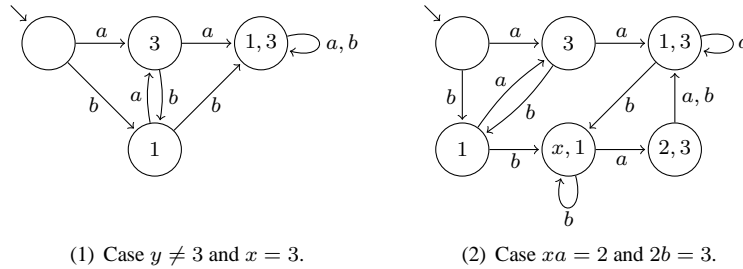
1. *if  $z = 1$ , then  $y = 3$  and it is  $xa \neq 2$  or  $2b \neq 3$ ,*
2. *if  $z = 2$  and  $3 \notin \{x, y\}$ , then  $1 \in \{x, y\}$  and  $3b \in \{x, y\}$ ,*
3. *if  $z = 2$  and  $3 \in \{x, y\}$ , if  $q \in \{x, y\} \setminus \{3\}$  then  $qa \neq 2$  or  $1b \neq y$ ,*
4. *if  $b = [x, y] \setminus z$ ,  $z \notin \{1, 2\}$ ,  $x \in \{1, 2\}$  and  $za \in \{1, 2\}$ , then  $y = 3$ .*

Moreover, if  $\mathcal{A}$  is proper and 3-compressible, then either the word  $b^2a^2$  or  $b^2ab^2$  or  $a^2b^2$  or  $a^2ba^2$  3-compresses  $\mathcal{A}$ .

**Proof:** First observe that from Corollary 24, a proper 3-compressible  $(4, 4)$ -automaton always satisfies the antecedent of one of the above conditions, so all the possible cases are taken into account.

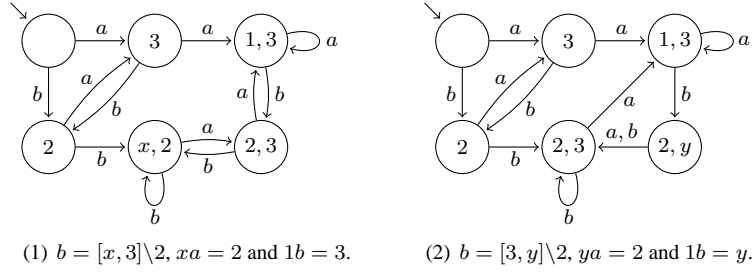
We start proving that a  $(4, 4)$ -automaton that does not satisfy conditions 1.-4. is not proper or it is not 3-compressible.

1. Let condition 1. be false, i.e.,  $b = [x, y] \setminus 1$  but either  $y \neq 3$  or  $xa = 2$  and  $2b = 3$ . Observe that if  $y \neq 3$  then from the previous corollary we have  $x = 3$ . The 3MSA in Fig. 37 proves that condition 1. is false, then  $\mathcal{A}$  is not 3-compressible.



**Fig. 37:** 3MSA for the case in which condition 1. of Proposition 25 is false.

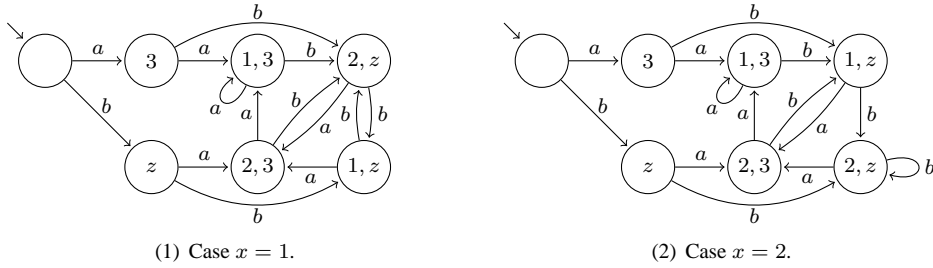
2. Let condition 2. be false, i.e.,  $b = [x, y] \setminus 2$  and  $3 \notin \{x, y\}$  but either  $1 \notin \{x, y\}$  or  $3b \notin \{x, y\}$ . From conditions 3. and 4. of Lemma 23 we have that in this cases  $\mathcal{A}$  is not proper, as it is 3-compressed either by  $ab^2$  or by  $a^2b$ .
3. Let condition 3. be false. If  $b = [x, 3] \setminus 2$  but  $xa = 2$  and  $1b = 3$ , then the 3MSA in Fig. 38(1) proves that  $\mathcal{A}$  is not 3-compressible. Else, if  $b = [3, y] \setminus 2$  but  $ya = 2$  and  $1b = y$ , then the 3MSA in Fig. 38(2) proves  $\mathcal{A}$  is not 3-compressible.



**Fig. 38:** 3MSA for the case in which condition 3. of Proposition 25 is false.

4. Let condition 4. be false, i.e.,  $b = [x, y] \setminus z$ ,  $z \notin \{1, 2\}$ ,  $x \in \{1, 2\}$  and  $za \in \{1, 2\}$ , but  $y \neq 3$ . We have to consider the following subcases.

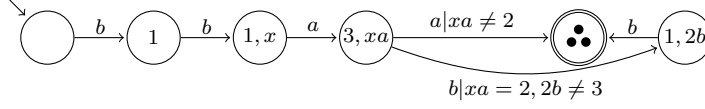
- (a) Let  $za = 1$ , hence  $z = 3$  and  $zb = x$ . If  $x = 1$ , then for all  $w \in \{a, b\}^+$  with  $|w| \geq 2$ , we have  $\mathcal{M}(w) = \{1, 3\}$ , and then  $\mathcal{A}$  is not 3-compressible. If  $x = 2$ , then by condition 4 of Lemma 23 we have  $y = 1$ , and for all  $w \in \{a, b\}^+$  we have  $M(wa) = \{1, 3\}$  and  $\mathcal{M}(wb) = \{2, 3\}$  and again  $\mathcal{A}$  is not 3-compressible.
- (b) Let  $za = 2$ , hence  $z \neq 3$ . We consider two further subcases:
  - i. if  $x = 1$ , then  $3b \neq 1$  and by condition 2. of Lemma 23  $za = y = 2$ , and by condition 3. of Lemma 23 we have  $3b = 2$ . The 3MSA in Fig. 39(1) proves that  $\mathcal{A}$  is not 3-compressible;
  - ii. if  $x = 2$ , then by condition 4. of Lemma 23 we have  $y = 1$  and by condition 3. of Lemma 23 we have  $3b = 1$ . The 3MSA in Fig. 39(2) proves that  $\mathcal{A}$  is not 3-compressible.



**Fig. 39:** 3MSA for the case in which condition 4. of Proposition 25 is false and  $za = 2$ .

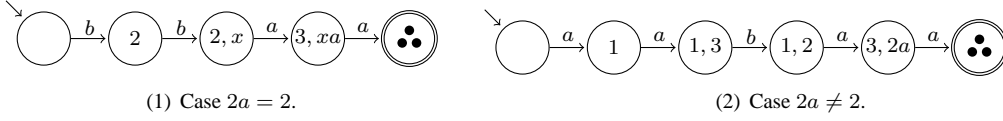
Conversely, we find for each automaton satisfying conditions 1.-4. a (short) 3-compressing word.

1. Let  $b = [x, 3] \setminus 1$  and  $xa \neq 2$  or  $2b \neq 3$ . Observe that  $xa \neq 1$ , as  $3a = 1$  and  $x \neq 3$ . The P3MSA in Fig. 40 proves that either the word  $b^2a^2$  or  $b^2ab^2$  3-compresses  $\mathcal{A}$ .
2. Let  $b = [x, y] \setminus 2$ ,  $3 \notin \{x, y\}$ ,  $1 \in \{x, y\}$  and  $3b \in \{x, y\}$ . There are two subcases.



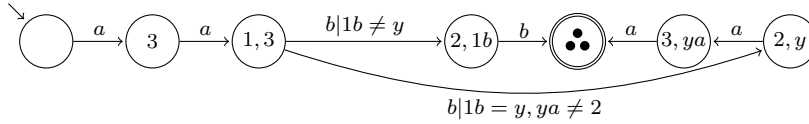
**Fig. 40:** P3MSA for the case  $b = [x, 3] \setminus 1$  and  $xa \neq 2$  or  $2b \neq 3$ .

- (a) Let  $x = 1$ , then  $b = [1, y] \setminus 2$ ,  $y \neq 3$ ,  $3b = y$  (as  $2b = 1$ ) and  $2a \neq 1$  (as  $3a = 1$ ). So  $\mathcal{M}(b^2a) = \{2a, 3\}$  and if  $2a \neq 2$ , then  $|\mathcal{M}(b^2a^2)| = 3$ , else if  $2a = 2$ , then  $|\mathcal{M}(b^2ab)| = 3$ , so either the word  $b^2a^2$  or  $b^2ab$  3-compresses  $\mathcal{A}$ .
- (b) Let  $y = 1$ , then  $b = [x, 1] \setminus 2$ ,  $x \neq 3$ ,  $3b = 1$  (as  $2b = x$ ) and  $2a \neq 1$  (as  $3a = 1$ ). If  $2a = 2$ , then the P3MSA in Fig. 41(1) proves that the word  $a^2b^2$  3-compresses  $\mathcal{A}$ . If  $2a \neq 2$ , then the P3MSA in Fig. 41(2) proves that the word  $a^2ba^2$  3-compresses  $\mathcal{A}$ .



**Fig. 41:** P3MSA for the case  $b = [x, 1] \setminus 2$ ,  $x \neq 3$  and  $3b \in \{x, 1\}$ .

3. If  $b = [x, 3] \setminus 2$  and  $xa \neq 2$  or  $1b \neq 3$ , then observe that  $xa \neq 1$ , as  $3a = 1$  and  $x \neq 3$ . If  $xa \neq 2$ , then  $\mathcal{M}(b^2) = \{2, x\}$ ,  $\mathcal{M}(b^2a) = \{3, xa\}$  and  $|\mathcal{M}(b^2a^2)| = 3$ . If  $xa = 2$  and  $1b \neq 3$ , then  $\mathcal{M}(a^2) = \{1, 3\}$ ,  $\mathcal{M}(a^2b) = \{1b, 2\}$  and  $|\mathcal{M}(a^2b^2)| = 3$ . Then either the word  $b^2a^2$  or  $a^2b^2$  3-compresses  $\mathcal{A}$ . If  $b = [3, y] \setminus 2$  and  $ya \neq 2$  or  $1b \neq y$ , then observe that  $ya \neq 1$ , as  $3a = 1$  and  $y \neq 3$ . The P3MSA in Fig. 42 proves that either the word  $a^2b^2$  or  $a^2ba^2$  3-compresses  $\mathcal{A}$ .



**Fig. 42:** P3MSA for the case  $b = [3, y] \setminus 2$  and  $ya \neq 2$  or  $1b \neq y$ .

4. Let  $b = [x, 3] \setminus z$ ,  $z \notin \{1, 2\}$ ,  $x \in \{1, 2\}$  and  $za \in \{1, 2\}$ . Observe that  $z \neq 3$  and  $3a = 1$ , so  $za \neq 1$ , and then it is always  $za = 2$ . We have to consider two subcases.
- (a) Let  $x = 1$ , then  $b = [1, 3] \setminus z$ ,  $zb = 1$ ,  $3b \neq 1$  and  $\mathcal{M}(a^2b) = \{z, 3b\}$ . If  $3b \neq 2$ , then  $\mathcal{M}(a^2ba) = \{za, 3ba, 3\}$  and  $a^2ba$  3-compresses  $\mathcal{A}$ . Else, if  $3b = 2$ , then  $\mathcal{M}(a^2b^2) = \{1, z, 2b\}$  and  $a^2b^2$  3-compresses  $\mathcal{A}$ .
- (b) Let  $x = 2$ , then  $b = [2, 3] \setminus z$ ,  $3b \neq 2$  (as  $zb = 2$ ) and  $\mathcal{M}(a^2b) = \{z, 1b\}$ . If  $1b \neq 3$ , then  $\mathcal{M}(a^2b^2) = \{2, 1b^2, z\}$  and  $a^2b^2$  3-compresses  $\mathcal{A}$ . Else, if  $1b = 3$ , then  $\mathcal{M}(a^2ba) = \{za, 1ba, 3\}$  and  $a^2ba$  3-compresses  $\mathcal{A}$ .

□



## 5 Finding lower and upper bounds for $c(3, 2)$

Collecting the words arising from the previous propositions, and taking into account that the roles of letters  $a$  and  $b$  are interchangeable, then each 3-full word, containing as factors the words in the set  $W$ :

$$W = \{ab^2ab^2a, ab^2a^2b^2a, abab^2aba, ab^3aba, abab^3a, ab^3ab^3a, ba^2baba^2b, a^2b^3a, ba^2ba^2b, ba^2b^2a^2b, baba^2bab, ba^3bab, baba^3b, ba^3ba^3b, ab^2abab^2a, b^2a^3b, a^2b^3a^2, b^2a^3b^2\}$$

is a 3-collapsing word on a two letter alphabet. Remark that  $a^2b^3a$  and  $b^2a^3b$  are factors of  $a^2b^3a^2$  and  $b^2a^3b^2$  respectively, the reason for which they occur in  $W$  will be clear in the sequel.

Then, in order to construct a short 3-collapsing word we need to find a word having as factor all the words above, *i.e.*, to solve the Shortest Common Supersequence problem (SCS) for  $W$ . It is well-known that SCS is NP-complete (Raiha and Ukkonen (1981)) also over a two letter alphabet, thus approximation algorithms are often used. Nevertheless, the cost of finding a good approximation is comparable to the cost of finding an optimal solution, as in Karpinski and Schmied (2013) the authors prove that the problem to approximate within any constant approximation factor better than 333/332 is NP-hard. On the other hand, efficient algorithms give poor approximation, as the best one yields an approximation factor of  $\approx 2$ , *i.e.*, it produces a solution whose length is about twice that of the optimum (Karpinski and Schmied (2013); Turner (1989)).

So, as no near-optimal solutions can be found in reasonable time, we decide to code the problem in the bounded satisfiability problem for a set of linear-time temporal logic (LTL) formulae, for which we developed a tool (Bersani et al. (2014)). More precisely, let  $S$  be a propositional letter, a word  $w$  of length  $n$  is coded in the LTL formula  $\mathbf{w}$  such that it is satisfied if and only if for all  $1 \leq i \leq n$  the  $i$ -th letter of  $w$  is “a” if and only if at the  $i$ -th time instant  $S$  is true. E.g., the word  $aba$  is encoded in the formula  $S \wedge (\mathbf{X}(\neg S \wedge \mathbf{X}(S)))$ , where  $\mathbf{X}$  is the “next” operator. Then we look for the shortest model that satisfy the formula  $\bigwedge_{w \in W} \mathbf{w}$ , obtaining that the shortest word having as factor all the words in  $W$  has length 55.

However, we were well aware that such “greedy approach”, *i.e.*, to find an optimal solution for each subcase and combining them to obtain a global solution, is not suitable in order to achieve a global optimum.

Then we observed that the words  $a^2b^3a^2$  and  $b^2a^3b^2$  are needed only to solve a special subcase of  $(3, 4)$ -automata, so we try to replace them with longest factor in order to obtain a shorter 3-collapsing word.

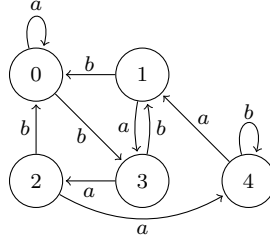
Actually, the shortest word having as factor the words in  $W \setminus \{a^2b^3a^2, b^2a^3b^2\}$  is

$$w_3 = b^2a^3b \underbrace{a^3b^3aba^2}_v baba^2ba^2b^2a^2b^2ab^2aba \underbrace{b^2aba^3bab^3}_u ab^3a$$

has as factor the word  $u = b^2aba^3bab^2$  and  $v = a^2b^3aba^2$ . As  $u$  belongs to the language  $\mathcal{L}$  defined in Proposition 22, case 2.b.iii, and  $v$  belongs to the dual of  $\mathcal{L}$ , this proves that  $w_3$  is a 3-collapsing word of length 53.

It is known that in general the language  $\mathcal{C}_{k,t}$  of  $k$ -collapsing words on an alphabet of  $t$  letters differs from the language  $\mathcal{S}_{k,t}$  of  $k$ -synchronizing words on the same alphabet. However, this not excludes that in some cases  $c(k, t) = s(k, t)$ . Up to now it was only known that  $c(2, 2) = s(2, 2)$  (Sauer and Stone (1991)) and that  $c(2, 3) \neq s(2, 3)$  (Ananichev and Petrov (2003)). We find a counterexample proving that  $c(3, 2) \neq s(3, 2)$  (and so  $c(3, 2) \geq 34$ ). In fact, the automaton in Fig. 43 is 3-compressible (and also 3-synchronizing), but the word  $s_{3,2}$  do not compresses it. On the other hand, its dual  $\bar{s}_{3,2}$  synchronizes

it. We check with a computer calculation that any 3-compressible 5-states automaton over a two letter alphabet, is 3-compressed either by  $s_{3,2}$  or by  $\bar{s}_{3,2}$ , so we believe interesting to understand if it is always the case for every 3-compressible automata.



**Fig. 43:** A semiautomaton which is not 3-compressed by  $s_{3,2}$ :  $Q_{\bar{s}_{3,2}} = \{3\}$ ,  $Q_{s_{3,2}} = \{0, 1, 3\}$ .

## 6 Conclusion

Although very technical, our analysis can be effectively exploited in order to obtain more general results and to investigate some conjectures. In Cherubini and Kisielewicz (2014, 2015), the authors exploit the characterization of  $(\mathbf{3}, \mathbf{p})$ -automata (Proposition 10) to prove that the problem of recognizing whether a binary word is 3-collapsing is co-NP-complete.

Moreover, the word  $w_3$  can be used to improve the procedure arising from Margolis et al. (2004) (Theorem 3.5) to obtain shorter  $k$ -collapsing words for  $k \geq 4$ . In particular it follows that  $c(4, 2) \leq 1741$  and  $c(5, 2) \leq 109941$ . Though very lengthy, they can be effectively used in testing the compressibility of an automaton. In particular this can accelerate the algorithm presented in Ananichev and Petrov (2003); Petrov (2008) to find short (possibly shortest) 4- and 5-synchronizing words.

## References

- D. S. Ananichev and I. V. Petrov. Quest for short synchronizing words and short collapsing words. In *WORDS. Proc. 4th Int. Conf.*, pages 411–418, 2003.
- D. S. Ananichev, I. V. Petrov, and M. V. Volkov. Collapsing words: A progress report. In *Developments in Language Theory*, volume 3572 of *LNCS*, pages 11–21, 2005.
- M. Bersani, A. Frigeri, A. Morzenti, M. Pradella, M. Rossi, and P. San Pietro. Constraint LTL satisfiability checking without automata. *Journal of Applied Logic*, 12(4):522 – 557, 2014.
- A. Cherubini. Synchronizing and collapsing words. *Milan J. Math.*, 75:305–321, 2007.
- A. Cherubini and A. Kisielewicz. Binary 3-compressible automata. In *Proceedings of the 15th Italian Conference on Theoretical Computer Science (ICTCS 2014)*, volume 1231 of *CEUR Workshop Proceedings*, pages 109–120, 2014.
- A. Cherubini and A. Kisielewicz. Recognizing 3-collapsing words over a binary alphabet. submitted, 2015.

- A. Cherubini, A. Kisielewicz, and B. Piochi. A bound for the length of shortest 2 collapsing words. *DMTCS*, 11(1):33–44, 2009.
- A. Cherubini, A. Frigeri, and B. Piochi. *Short 3-collapsing words over a 2-letter alphabet*, volume 6795 of *Lecture Notes in Computer Science*. 2011.
- M. Karpinski and R. Schmied. Improved inapproximability results for the shortest superstring and related problems. In *19th Computing: The Australasian Theory Symposium (CATS 2013), Adelaide, Australia*, volume 141 of *CRPIT*, pages 27–36, 2013.
- R. Kudlacik, A. Roman, and H. Wagner. Effective synchronizing algorithms. *Expert Syst. App.*, 39(14): 11746–11757, 2012.
- S. W. Margolis, J.-E. Pin, and M. V. Volkov. Words guaranteeing minimum image. *Internat. J. Foundations Comp. Sci.*, 15:259–276, 2004.
- I. Petrov. An algorithm for recognition of  $n$ -collapsing words. *Theoret. Comput. Sci.*, 391(1-2):99–108, 2008.
- R. Pöschel, M. V. Sapir, N. W. Sauer, M. G. Stone, and M. V. Volkov. Identities in full transformation semigroups. *Algebra Universalis*, 31:580–588, 1994.
- E. V. Pribavkina. On some properties of the language of 2-collapsing words. In *Developments in Language Theory*, volume 3572 of *LNCS*, pages 374–384, 2005.
- K. Raiha and E. Ukkonen. The shortest common supersequence problem over binary alphabet in NP-complete. *Theoretical Computer Science*, 16:187–198, 1981.
- N. Sauer and M. G. Stone. Composing functions to reduce image size. *Ars Combinatoria*, 1:171–176, 1991.
- A. N. Trahtman. An efficient algorithm finds noticeable trends and examples concerning the Černý conjecture. In *Proceedings of MFCS*, volume 4162 of *LNCS*, pages 789–800, 2006.
- J. Turner. Approximation algorithms for the shortest common superstring problem. *Information and Computation*, 83:1–20, 1989.
- M. V. Volkov. Synchronizing automata and the Černý conjecture. In *Language and automata theory and applications*, volume 5196, pages 11–27. *LNCS*, 2008.